# 3-Receiver Broadcast Channels with Common [1] and Confidential Messages

Yeow-Khiang Chia and Abbas El Gamal
Department of Electrical Engineering
Stanford University
Stanford, CA 94305, USA
Email: ykchia@stanford.edu, abbas@ee.stanford.edu

### Abstract

This paper establishes inner bounds on the secrecy capacity regions for the general 3-receiver broadcast channel with one common and one confidential message sets. We consider two setups. The first is when the confidential message is to be sent to two receivers and kept secret from the third receiver. Achievability is established using indirect decoding, Wyner wiretap channel coding, and the new idea of generating secrecy from a publicly available superposition codebook. The inner bound is shown to be tight for a class of reversely degraded broadcast channels and when both legitimate receivers are less noisy than the third receiver. The second setup investigated in this paper is when the confidential message is to be sent to one receiver and kept secret from the other two receivers. Achievability in this case follows from Wyner wiretap channel coding and indirect decoding. This inner bound is also shown to be tight for several special cases.

## I. Introduction

The wiretap channel was first introduced in the seminal paper by Wyner [1]. He considered a 2-receiver broadcast channel where sender $X$ wishes to communicate a message to receiver $Y$ while keeping it secret from the other receiver (eavesdropper) $Z$. Wyner showed that the secrecy capacity when the channel to the eavesdropper is a degraded version of the channel to the legitimate receiver is

$$C_{\mathrm{s}} = \max_{p(x)} (I(X;Y) - I(X;Z)).$$

The main coding idea is to randomly generate $2^{n(I(X;Y))}$ $x^n$ sequences and partition them into $2^{nR}$ message bins, where $R < I(X;Y) - I(X;Z)$. To send a message, a sequence from the message bin is *randomly* selected and transmitted. The legitimate receiver uniquely decodes the codeword and hence the message with high probability, while the message is kept asymptotically secret from the eavesdropper provided $R < C_{\mathrm{S}}$.

This result was extended by Csiszár and Körner [2] to general (non-degraded) 2-receiver broadcast channels with common and confidential messages. They established the secrecy capacity region, which is the optimal tradeoff between the common and private message rates and the eavesdropper's private message equivocation rate. In the special case of no common message, their result yields the secrecy capacity for the general wiretap channel,

$$C_{\mathrm{s}} = \max_{p(v)p(x|v)} (I(V;Y) - I(V;Z)).$$

The achievability idea is to use Wyner's wiretap channel coding for the channel from $V$ to $Y$ by randomly selecting a $v^n$ codeword from the message bin and then sending a random sequence $X^n$ generated according to $\prod_{i=1}^n p_{X|V}(x_i|v_i)$.

The work in [2] has been extended in several directions by considering different message demands and secrecy scenarios, e.g., see [3], [4]. However, with some notable exceptions such as [5] and [6], extending the result of Csiszár and Körner to general discrete memoryless broadcast channels with more than two receivers has remained open, since even the capacity region without secrecy constraints for the 3-receiver broadcast channel with degraded message sets is not known in general. The secrecy setup for the 3-receiver broadcast channel also has close connections to the compound wiretap channel model (see [7, Chapter 3] and references therein). Recently, Nair and El Gamal [8] showed that the straightforward extension of the Körner–Marton capacity region for the 2-receiver broadcast

channel with degraded message sets to more than 3 receivers is not optimal. They established an achievable rate region for the general 3-receiver broadcast channel and showed that it can be strictly larger than the straightforward extension of the Körner–Marton region.

In this paper, which is a much expanded version of [9], we establish inner and outer bounds on the secrecy capacity region for the 3-receivers broadcast channel with common and confidential messages. We consider two setups.

- *2-receiver, 1-eavesdropper*: Here the confidential message is to be sent to two receivers and kept secret from the third receiver (eavesdropper).
- *1-receiver, 2-eavesdroppers*: In this setup the confidential message is to be sent to one receiver and kept secret from the other two receivers.

To illustrate the main coding idea in our new inner bound for the 2-receiver, 1-eavesdropper setup, consider the special case where a message $M \in [1 : 2^{nR}]$ is to be sent reliably to receivers $Y_1$ and $Y_2$ and kept asymptotically secret from eavesdropper $Z$. A straightforward extension of the Csiszár–Körner [2] result for the 2-receiver wiretap channel yields the lower bound on the secrecy capacity

$$C_{\mathrm{S}} \geq \max_{p(v)p(x|v)} \min\{I(V;Y_1) - I(V;Z), I(V;Y_2) - I(V;Z)\}. \tag{1}$$

Now, suppose $Z$ is a degraded version of $Y_1$, then from Wyner's wiretap result, we know that $(I(V;Y_1) - I(V;Z)) \leq (I(X;Y_1) - I(X;Z))$ for all $p(v,x)$. However, no such inequality holds in general for the second term under the minimum. As a special case of the inner bound in Theorem 1, we show that the rate obtained by replacing $V$ by $X$ only in the first term in (1) is achievable, that is, we establish the lower bound

$$C_{\mathrm{S}} \geq \max_{p(v)p(x|v)} \min\{I(X;Y_1) - I(X;Z), I(V;Y_2) - I(V;Z)\}. \tag{2}$$

To prove achievability of (2), we again randomly generate $2^{n(I(V;Y_2)-\delta)}$ $v^n$ sequences and partition them into $2^{nR}$ bins, where $R = (I(V;Y_2) - I(V;Z))$. For each $v^n$ sequence, we randomly and conditionally independently generate $2^{nI(X;Z|V)}$ $x^n$ sequences. The $v^n$ and $x^n$ sequences are revealed to all parties, including the eavesdropper. To send a message $m$, the encoder randomly chooses a $v^n$ sequence from bin $m$. It then *randomly* chooses an $x^n$ sequence from the codebook for the selected $v^n$ sequence (instead of randomly generating an $X^n$ sequence as in the Csiszár–Körner scheme) and transmits it. Receiver $Y_2$ decodes $v^n$ directly, while receiver $Y_1$ decodes $v^n$ *indirectly* through $x^n$ [8]. In Section III, we show through an example that this new lower bound can be strictly larger than the extended Csiszár–Körner lower bound. We then show in Theorem 1 that this lower bound can be generalized further via Marton coding.

The rest of the paper is organized as follows. In the next section we present needed definitions. In Section III, we provide an alternative proof of achievability for the Csiszár–Körner 2-receiver wiretap channel that uses superposition coding and random codeword selection instead of random generating of the transmitted codeword. This technique is used in subsequent sections to establish the new inner bounds for the 3-receiver setups. In Section IV, we present the inner bound for the 2-receiver, 1-eavesdropper case. We show that this lower bound is tight for the reversely degraded product broadcast channel and when the eavesdropper is less noisy than both legitimate receivers. In Section V, we present inner and outer bounds for the 1-receiver, 2-eavesdropper setup for 3-receiver multilevel broadcast channel [10]. We show that the bounds coincide in several special cases.

## II. DEFINITIONS AND PROBLEM SETUP

Consider a 3-receiver discrete memoryless broadcast channel with input alphabet $\mathcal{X}$, output alphabets $\mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Y}_3$ and conditional pmfs $p(y_1, y_2, y_3|x)$. We investigate the following two setups.

### A. 2-Receivers, 1-Eavesdropper

Here the confidential message is to be sent to receivers $Y_1$ and $Y_2$ and is to be kept secret from the eavesdropper $Y_3 = Z$. A $(2^{nR_0}, 2^{nR_1}, n)$ code for this scenario consists of: (i) two messages $(M_0, M_1)$ uniformly distributed over $[1 : 2^{nR_0}] \times [1 : 2^{nR_1}]$; (ii) an encoder that randomly generates a codeword $X^n(m_0, m_1)$ according to the conditional pmf $p(x^n|m_0, m_1)$; and (iii) 3 decoders; the first decoder assigns to each received sequence $y_1^n$ an

estimate $(\hat{M}_{01}, \hat{M}_{11}) \in [1 : 2^{nR_0}] \times [1 : 2^{nR_1}]$ or an error message, the second decoder assigns to each received sequence $y_2^n$ an estimate $(\hat{M}_{02}, \hat{M}_{12}) \in [1 : 2^{nR_0}] \times [1 : 2^{nR_1}]$ or an error message, and the third decoder assigns to each received sequence $z^n$ an estimate $\hat{M}_{03} \in [1 : 2^{nR_0}]$ or an error message. The probability of error for this scenario is defined as

$$P_{e1}^{(n)} = \mathrm{P}\left\{\hat{M}_{0j} \neq M_0 \text{for } j = 1, 2, 3 \text{or } \hat{M}_{1j} \neq M_1 \text{for } j = 1, 2\right\}.$$

The equivocation rate at receiver $Z$, which measures the amount of uncertainty receiver $Z$ has about message $M_1$, is defined as $H(M_1|Z^n)/n$.

A secrecy rate tuple $(R_0, R_1, R_e)$ is said to be achievable if

$$\lim_{n\to\infty} P_{e1}^{(n)} = 0, \text{and}$$

$$\liminf_{n\to\infty} \frac{1}{n} H(M_1|Z^n) \geq R_e.$$

The *secrecy capacity region* is the closure of the set of achievable rate tuples $(R_0, R_1, R_e)$.

For this setup, we also consider the special case of *asymptotic perfect secrecy*, where no common message is to be sent to $Z$ and a confidential message, $M \in [1 : 2^{nR}]$, is to be sent to $Y_1$ and $Y_2$ only. The probability of error is as defined above with $R_0 = 0$ and $R_1 = R$. A secrecy rate $R$ is said to be achievable if there exists a sequence of $(2^{nR}, n)$ codes such that

$$\lim_{n\to\infty} P_{e1}^{(n)} = 0, \text{and}$$

$$\liminf_{n\to\infty} \frac{1}{n} H(M|Z^n) \geq R.$$

The *secrecy capacity*, $C_S$, is the supremum of all achievable rates.

### B. 1-Receiver, 2-Eavesdroppers

In this setup, the confidential message is to be sent to receiver $Y_1$ and kept secret from eavesdroppers $Y_2 = Z_2$ and $Y_3 = Z_3$. A $(2^{nR_0}, 2^{nR_1}, n)$ code consists of the same message sets and encoding function as in the 2-receiver, 1-eavesdropper case. The first decoder assigns to each received sequence $y_1^n$ an estimate $(\hat{M}_{01}, \hat{M}_1) \in [1 : 2^{nR_0}] \times [1 : 2^{nR_1}]$ or an error message, the second decoder assigns to each received sequence $z_2^n$ an estimate $\hat{M}_{02} \in [1 : 2^{nR_0}]$ or an error message, and the third decoder assigns to each received sequence $z_3^n$ an estimate $\hat{M}_{03} \in [1 : 2^{nR_0}]$ or an error message. The probability of error is

$$P_{e2}^{(n)} = \mathrm{P}\{\hat{M}_{0j} \neq M_0 \text{for } j = 1, 2, 3 \text{or } \hat{M}_1 \neq M_1\}.$$

The equivocation rates at the two eavesdroppers are $H(M_1|Z_2^n)/n$ and $H(M_1|Z_3^n)/n$, respectively.

A secrecy rate tuple $(R_0, R_1, R_{e2}, R_{e3})$ is said to be achievable if

$$\lim_{n\to\infty} P_{e2}^{(n)} = 0,$$

$$\liminf_{n\to\infty} \frac{1}{n} H(M_1|Z_j^n) \geq R_{ej}, \ \ j = 2, 3.$$

The *secrecy capacity region* is the closure of the set of achievable rate tuples $(R_0, R_1, R_{e2}, R_{e3})$. For simplicity of presentation, we consider only the special class of multilevel broadcast channels [10].

## III. 2-RECEIVER WIRETAP CHANNEL

We first revisit the 2-receiver wiretap channel, where a confidential message is to be sent to the legitimate receiver $Y$ and kept secret from the eavesdropper $Z$. The secrecy capacity for this case is a special case of the secrecy capacity region for the broadcast channel with common and confidential messages established in [2].

*Proposition 1:* The secrecy capacity of the 2-receiver wiretap channel is

$$C_S = \max_{p(v,x)} (I(V; Y) - I(V; Z)).$$

3

In the following, we provide a new proof of achievability for this result in which the second randomization step in the original proof is replaced by a random codeword selection from a *public* superposition codebook. As we will see, this proof technique allows us to use indirect decoding to establish new inner bounds for the 3-receiver wiretap channels.

*Proof of Achievability for Proposition 1*:

Fix $p(v, x)$. Randomly and independently generate sequences $v^n(l_0)$, $l_0 \in [1 : 2^{n\tilde{R}}]$, each according to $\prod_{i=1}^{n} p_V(v_i)$. Partition the set $[1 : 2^{n\tilde{R}}]$ into $2^{nR}$ bins $\mathcal{B}(m) = [(m-1)2^{n(\tilde{R}-R)} + 1 : m2^{n(\tilde{R}-R)}]$, $m \in [1 : 2^{nR}]$. For each $l_0 \in [1 : 2^{n\tilde{R}}]$, randomly and conditionally independently generate sequences $x^n(l_0, l_1)$, $l_1 \in [1 : 2^{n\tilde{R}_1}]$, each according to $\prod_{i=1}^{n} p_{X|V}(x_i|v_i)$. The codebook $\{(v^n(l_0), x^n(l_0, l_1))\}$ is revealed to all parties. To send the message $m$, an index $L_0 \in \mathcal{B}(m)$ is selected uniformly at random (as in Wyner's original proof). The encoder then randomly and independently selects an index $L_1$ and transmits $x^n(L_0, L_1)$. Receiver $Y$ decodes $L_0$ by finding the unique index $\hat{l}_0$ such that $(v^n(\hat{l}_0), y^n) \in \mathcal{T}_\epsilon^{(n)}$. By the law of large numbers and the packing lemma [11, Chapter 3], the average probability of error approaches zero as $n \to \infty$ if $\tilde{R} < (V; Y) - \delta(\epsilon)$.

We now show that $I(M; Z^n|\mathcal{C}) \leq n\delta(\epsilon)$. Considering the mutual information between $Z^n$ and $M$, averaged over the random codebook $\mathcal{C}$, we have

$$
\begin{aligned}
I(M; Z^n|\mathcal{C}) &= I(M, L_0, L_1; Z^n|\mathcal{C}) - I(L_0, L_1; Z^n|M, \mathcal{C}) \\
&\overset{(a)}{\leq} I(X^n; Z^n|\mathcal{C}) - H(L_0, L_1|M, \mathcal{C}) + H(L_0, L_1|M, Z^n, \mathcal{C}) \\
&\leq \sum_{i=1}^{n} I(X_i; Z_i|\mathcal{C}) - n(\tilde{R} - R) - n\tilde{R}_1 + H(L_0, L_1|M, Z^n, \mathcal{C}) \\
&\leq nI(X; Z) - n(\tilde{R} + \tilde{R}_1 - R) + H(L_0|M, Z^n, \mathcal{C}) + H(L_1|L_0, Z^n, \mathcal{C}).
\end{aligned}
\tag{3}
$$

$(a)$ follows since $(M, L_0, L_1, \mathcal{C}) \to X^n \to Z^n$ from the discrete memoryless property of the channel. The last step follows from follows since $H(Z_i|\mathcal{C}) \leq H(Z_i) = H(Z)$ and $H(Z_i|X_i, \mathcal{C}) = \sum_{\mathcal{C}} p(\mathcal{C})p(x_i|\mathcal{C})H(Z|x_i, \mathcal{C}) = \sum_{\mathcal{C}} p(\mathcal{C})p(v_i|\mathcal{C})H(Z|x_i) = H(Z|X)$ It remains to upper bound $H(L_0|M, Z^n, \mathcal{C})$ and $H(L_1|L_0, Z^n, \mathcal{C})$. By symmetry of codebook construction, we have

$$
\begin{aligned}
H(L_0|M, Z^n, \mathcal{C}) &= 2^{-nR} \sum_{m=1}^{2^{nR}} H(L_0|M = m, Z^n, \mathcal{C}) \\
&= H(L_0|Z^n, M = 1, \mathcal{C}), \\
H(L_1|L_0, Z^n, \mathcal{C}) &= 2^{-n\tilde{R}} \sum_{l_0} H(L_1|L_0 = l_0, Z^n, \mathcal{C}) \\
&= H(L_1|L_0 = 1, v^n(1), Z^n, \mathcal{C}).
\end{aligned}
$$

To further bound these terms, we use the following key lemma.

*Lemma 1:* Let $(U, V, Z) \sim p(u, v, z)$, $S \geq 0$ and $\epsilon > 0$. Let $U^n$ be a random sequence distributed according to $\prod_{i=1}^{n} p_U(u_i)$. Let $V^n(l)$, $l \in [1 : 2^{nS}]$, be a set of random sequences that are conditionally independent given $U^n$ and each distributed according to $\prod_{i=1}^{n} p_{V|U}(v_i|u_i)$, and let $\mathcal{C} = \{U^n, V^n(l)\}$. Let $L \in [1 : 2^{nS}]$ be a random index with an arbitrary probability mass. Then, if $\mathrm{P}\{(U^n, V^n(L), Z^n) \in \mathcal{T}_\epsilon^{(n)}\} \to 1$ as $n \to \infty$ and $S \geq I(V; Z|U) + \delta(\epsilon)$, there exists a $\delta'(\epsilon) > 0$, where $\delta'(\epsilon) \to 0$ as $\epsilon \to 0$, such that, for $n$ sufficiently large,

$$
H(L|Z^n, U^n, \mathcal{C}) \leq n(S - I(V; Z|U)) + n\delta'(\epsilon).
$$

The proof of this lemma is given in Appendix I. An illustration of the random sequence structure is given in Figure 1.

Now, returning to (3), we note that $\mathrm{P}\{(V^n(L_0), X^n(L_0, L_1), Z^n) \in \mathcal{T}_\epsilon^{(n)}\} \to 1$ as $n \to \infty$ by law of large numbers. Hence, we can apply Lemma 1 to obtain

$$
H(L_0|Z^n, M = 1, \mathcal{C}) \leq n((\tilde{R} - R) - I(V; Z)) + n\delta(\epsilon),
\tag{4}
$$

$$
H(L_1|L_0 = 1, V^n, Z^n, \mathcal{C}) \leq n(\tilde{R}_1 - I(X; Z|V)) + n\delta(\epsilon),
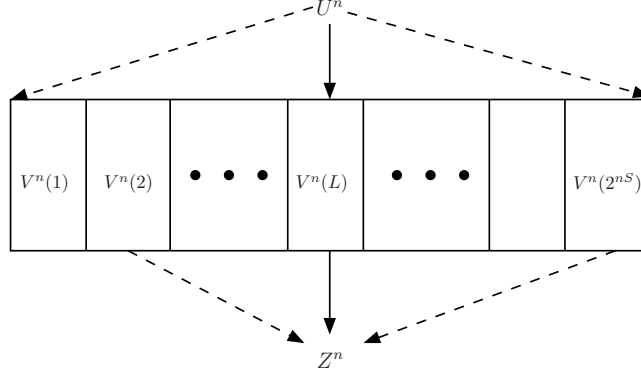\tag{5}
$$

4

Fig. 1: Structure of random sequences in Lemma 1. $V^n(l)$ is generated according to $\prod_{i=1}^{n} p_{V|U}(v_i|u_i)$. Solid arrows represent the sequence pair $(U^n, V^n(L), Z^n)$ while the dotted arrows to $Z^n$ represent the other $V^n$ sequences jointly typical with the $(U^n, Z^n)$ pair. Lemma 1 gives an upper bound on the number of $V^n$ sequences that can be jointly typical with a $(U^n, Z^n)$ pair.

if $\tilde{R} - R \geq I(V; Z) + \delta(\epsilon)$ and $\tilde{R}_1 \geq I(X; Z|V) + \delta(\epsilon)$. Substituting from inequalities (4) and (5) into (3) shows that $I(M; Z^n|\mathcal{C}) \leq 2n\delta(\epsilon)$. We then recover the original asymptotic secrecy rate by noting that the constraint of $\tilde{R}_1 \geq I(X; Z|V)$ is not tight. This completes the proof of Proposition 1.

*Remark 3.1*: In the proof of Proposition 1 in [2], the encoder transmits a randomly generated codeword $X^n \sim \prod_{i=1}^{n} p_{X|V}(x_i|v_i)$. Although replacing random $X^n$ generation by superposition coding and random codeword selection in our alternative proof does not increase the achievable secrecy rate for the 2-receiver wiretap channel, it can increase the rate when there are more than one legitimate receiver, as we show in the next sections.

## IV. 2-RECEIVERS, 1-EAVESDROPPER WIRETAP CHANNEL

We establish an inner bound on the secrecy capacity for the 3-receiver wiretap channel with one common and one confidential message when the confidential message is to be sent to receivers $Y_1$ and $Y_2$ and kept secret from receiver $Z$. In the following subsection, we consider the case where $M_0 = \emptyset$ and $M_1 = M \in [1 : 2^{nR}]$ is to be kept asymptotically secret from $Z$. This result is then extended in Subsection IV-B to establish an inner bound on the secrecy capacity region.

### A. Asymptotic perfect secrecy

We establish the following lower bound on secrecy capacity for the case where a confidential message is to be sent to receivers $Y_1$ and $Y_2$ and kept secret from the eavesdropper $Z$.

*Theorem 1:* The secrecy capacity of the 2-receiver, 1-eavesdropper setup with one confidential message and asymptotic secrecy is lower bounded as follows

$$C_S \geq \min\{I(V_0, V_1; Y_1|Q) - I(V_0, V_1; Z|Q), I(V_0, V_2; Y_2|Q) - I(V_0, V_2; Z|Q)\}$$

for some $p(q, v_0, v_1, v_2, x) = p(q, v_0)p(v_1, v_2|v_0)p(x|v_1, v_2, v_0)$ such that $I(V_1, V_2; Z|V_0) \leq I(V_1; Z|V_0) + I(V_2; Z|V_0) - I(V_1; V_2|V_0)$.
In addition to superposition coding and the new coding idea discussed in the previous section, Theorem 1 also uses Marton coding [12].

For clarity, we first establish the following Corollary 1.

*Corollary 1:* The secrecy capacity for the 2-receiver, 1-eavesdropper with one confidential message and asymptotic secrecy is lower bounded as follows

$$C_S \geq \max_{p(q)p(v|q)p(x|v)} \min\{I(X; Y_1|Q) - I(X; Z|Q), I(V; Y_2|Q) - I(V; Z|Q)\}.$$

5

*Remark 4.1*: Consider the case where $X \to Y_1 \to Z$ form a Markov chain. Then, we can show that Theorem 1 reduces to Corollary 1, i.e., the achievable secrecy rate is not increased by using Marton coding when $X \to Y_1 \to Z$ (or $X \to Y_2 \to Z$ by symmetry) form a Markov chain. To see this, note that $(I(X;Y_1|Q) - I(X;Z|Q)) \geq (I(V_1,V_0;Y_1|Q) - I(V_1,V_0;Z|Q))$ for all $V_1$ if $X \to Y_1 \to Z$. Next, note that we can set $V = (V_0,V_2)$ in Corollary 1 to obtain the rate in Theorem 1.

*Proof of Corollary 1*:

*Codebook generation*: Randomly and independently generate the time-sharing sequence $q^n$ according to $\prod_{i=1}^n p_Q(q_i)$. Next, randomly and conditionally independently generate $2^{n\tilde{R}}$ sequences $v^n(l_0)$, $l_0 \in [1 : 2^{n\tilde{R}}]$, each according to $\prod_{i=1}^n p_{V|Q}(v_i|q_i)$. Partition the set $[1 : 2^{n\tilde{R}}]$ into $2^{nR}$ equal size bins $\mathcal{B}(m)$, $m \in [1 : 2^{nR}]$. For each $l_0$, conditionally independently generate sequences $x^n(l_0,l_1)$, $l_1 \in [1 : 2^{n\tilde{R}_1}]$, each according to $\prod_{i=1}^n p_{X|V}(x_i|v_i)$.

*Encoding*: To send a message $m \in [1 : 2^{nR}]$, randomly and independently choose an index $L_0 \in \mathcal{C}(m)$ and an index $L_1 \in [1 : 2^{n\tilde{R}_1}]$, and send $x^n(L_0,L_1)$.

*Decoding*: Assume without loss of generality that $L_0 = 1$ and $m = 1$. Receiver $Y_2$ finds $L_0$, and hence $m$, via joint typicality decoding. By the law of large number and the packing lemma, the probability of error approaches zero as $n \to \infty$ if

$$\tilde{R} < I(V;Y_2|Q) - \delta(\epsilon).$$

Receiver $Y_1$ finds $L_0$ (and hence $m$) via indirect decoding. That is, it declares that $\hat{L}_0$ is sent if it is the unique index such that $(q^n, v^n(\hat{L}_0), x^n(\hat{L}_0,l_1), Y_1^n) \in \mathcal{T}_\epsilon^{(n)}$ for some $l_1 \in [1 : 2^{n\tilde{R}_1}]$. To analyze the average probability of error $P(\mathcal{E})$, define the error events

$$\mathcal{E}_{10} = \{(Q^n, X^n(1,1), Y_1^n) \notin \mathcal{T}_\epsilon^{(n)}\},$$
$$\mathcal{E}_{11} = \{(Q^n, X^n(l_0,l_1), Y_1^n) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } l_0 \neq 1\}.$$

Then, by union of events bound the probability of error is upper bounded as

$$P(\mathcal{E}) \leq P\{\mathcal{E}_{10}\} + P\{\mathcal{E}_{11}\}.$$

Now by law of large numbers, $P\{\mathcal{E}_{10}\} \to 0$ as $n \to \infty$. Next consider

$$\begin{aligned}
P\{\mathcal{E}_{11}\} &\leq \sum_{l_0 \neq 1} \sum_{l_1} P\{(Q^n, V^n(l_0), X^n(l_0,l_1), Y_1^n) \in \mathcal{T}_\epsilon^{(n)}\} \\
&\leq \sum_{l_0 \neq 1} \sum_{l_1} 2^{-n(I(V,X;Y_1|Q)-\delta(\epsilon))} \\
&\leq 2^{n(\tilde{R}+\tilde{R}_1-I(V,X;Y_1|Q)+\delta(\epsilon))}.
\end{aligned}$$

Hence, $P\{\mathcal{E}_{11}\} \to 0$ as $n \to \infty$ if

$$\tilde{R} + \tilde{R}_1 < I(X;Y_1|Q) - \delta(\epsilon).$$

*Analysis of equivocation rate*: To bound the equivocation rate term $H(M|Z^n, \mathcal{C})$, we proceed as before and show that the $I(M;Z^n|\mathcal{C}) \leq 2n\delta(\epsilon)$. Note that the only difference between this case and the analysis for the 2-receiver case in Section II is the addition of the time-sharing random variable $Q$. Since $P\{(Q^n, V^n(L_0), X^n(L_0,L_1), Z^n) \in \mathcal{T}_\epsilon^{(n)}\} \to 1$ as $n \to \infty$, we can apply Lemma 1 (with the addition of the time sharing random variable). Following the analysis in Section II, it is easy to see that $I(M;Z^n|\mathcal{C}) \leq 2n\delta(\epsilon)$ if

$$\tilde{R} - R \geq I(V;Z|Q) + \delta(\epsilon),$$
$$\tilde{R}_1 \geq I(X;Z|V) + \delta(\epsilon).$$

Finally, using Fourier–Motzkin elimination on the set of inequalities completes the proof of achievability.

Before proving Theorem 1, we show through an example that the lower bound in Corollary 1 can be strictly larger than the rate of the straightforward extension of the Csiszár–Körner scheme to the 2-receiver, 1-eavesdropper setting,

$$R_{CK} = \max_{p(q)p(v|q)p(x|v)} \min\{I(V; Y_1|Q) - I(V; Z|Q), I(V; Y_2|Q) - I(V; Z|Q)\}.. \tag{6}$$

Note that Theorem 1 includes $R_{CK}$ as a special case (through setting $V_0 = V_1 = V_2 = V$ in Theorem 1).

*Example*: Consider the multilevel product broadcast channel example [8] in Figure 2, where $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{Y}_{12} = \mathcal{Y}_{21} = \{0, 1\}$, and $\mathcal{Y}_{11} = \mathcal{Z}_1 = \mathcal{Z}_2 = \{0, E, 1\}$. The channel conditional probabilities are specified in Figure 2.
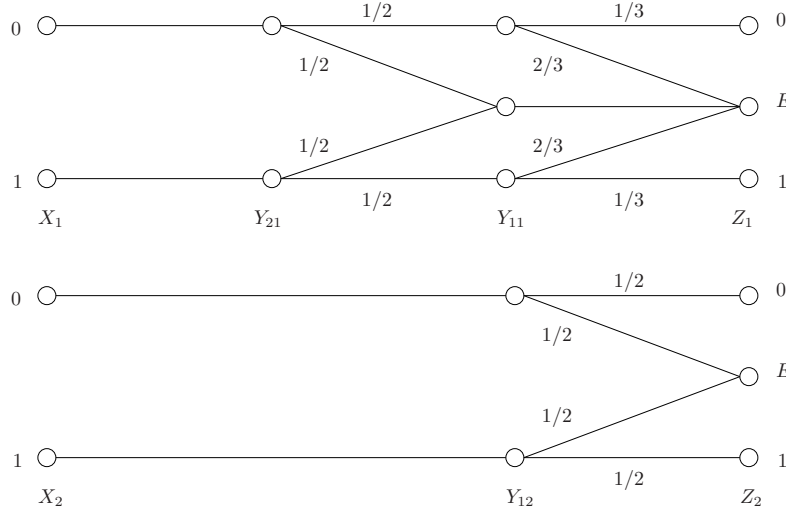


Fig. 2: Multilevel broadcast channel

In Appendix II, we show that $R_{CK} < 5/6$. In contrast, using Corollary 1, we can achieve a rate of $5/6$, which shows that the rate given in Theorem 1 can be strictly larger than using the straightforward extension of the Csiszár–Körner scheme.

We now turn to the proof of Theorem 1, which utilizes Marton coding in addition to the ideas already introduced.

*Proof of Theorem 1*:

*Codebook generation*: Randomly and independently generate a time-sharing sequence $q^n$ according to $\prod_{i=1}^{n} p_Q(q_i)$. Randomly and conditionally independently generate sequences $v_0^n(l_0)$, $l_0 \in [1 : 2^{n\tilde{R}}]$, each according to $\prod_{i=1}^{n} p_{V_0|Q}(v_{0i}|q_i)$. Partition the set $[1 : 2^{n\tilde{R}}]$ into $2^{nR}$ bins, $\mathcal{B}(m)$, $m \in [1 : 2^{nR}]$ as before. For each $l_0$, randomly and conditionally independently generate sequences $v_1^n(l_0, t_1)$, $t_1 \in [1 : 2^{nT_1}]$, each according to $\prod_{i=1}^{n} p_{V_1|V_0}(v_{1i}|v_{0i})$. Partition the set $[1 : 2^{nT_1}]$ into $2^{n\tilde{R}_1}$ equal size bins, $\mathcal{B}(l_0, l_1)$. Similarly, for each $l_0$, generate sequences $v_2^n(l_0, t_2)$, $t_2 \in [1 : 2^{nT_2}]$, each according to $\prod_{i=1}^{n} p_{V_2|V_0}(v_{2i}|v_{0i})$, and partition $[1 : 2^{nT_1}]$ into $2^{n\tilde{R}_2}$ equal size bins, $\mathcal{B}(l_0, l_2)$. Finally, for each product bin $\mathcal{B}(l_0, l_1) \times \mathcal{B}(l_0, l_2)$, find a jointly typical sequence pair $(v_1^n(l_0, t_1(l_0, l_1)), v_2^n(l_0, t_2(l_0, l_2)))$. If there is more than one such pair, randomly and uniformly pick one of them. This encoding step succeeds with probability of error that approaches zero as $n \to \infty$, if [13]

$$\tilde{R}_1 + \tilde{R}_2 < T_1 + T_2 - I(V_1; V_2|V_0) - \delta(\epsilon).$$

*Encoding*: To send message $m$, the encoder first randomly chooses an index $L_0 \in \mathcal{B}(m)$. It then randomly chooses a product bin indices $(L_1, L_2)$ and selects the jointly typical sequence pair $(v_1^n(L_0, t_1(L_0, L_1)), v_2^n(L_0, t_2(L_0, L_2)))$. Finally, the encoder generates a codeword $X^n$ at random according to $\prod_{i=1}^{n} p_{X|V_0, V_1, V_2}(x_i|v_{0i}, v_{1i}, v_{2i})$ and transmits it.

*Decoding and analysis of the probability of error*: Receiver $Y_1$ decodes $L_0$ and hence $m$ indirectly by finding the unique index $\hat{l}_0$ such that $(v_0^n(\hat{l}_0), v_1^n(\hat{l}_0, t_1), y_1^n) \in \mathcal{T}_\epsilon^{(n)}$ for some $t_1 \in [1 : 2^{nT_1}]$. Similarly, receiver $Y_2$

finds $L_0$ (and hence $m$) indirectly by finding the unique index $\hat{l}_0$ such that $(v_0^n(\hat{l}_0), v_2^n(\hat{l}_0, T_2)) \in \mathcal{T}_\epsilon^{(n)}$ for some $l_2 \in [1 : 2^{nT_2}]$. Following the analysis given earlier, it is easy to see that these steps succeed with probability of error that approaches zero as $n \to \infty$ if

$$\tilde{R} + T_1 < I(V_0, V_1; Y_1 | Q) - \delta(\epsilon),$$
$$\tilde{R} + T_2 < I(V_0, V_1; Y_2 | Q) - \delta(\epsilon).$$

*Analysis of equivocation rate*: A codebook $\mathcal{C}$ induces a joint pmf on $(M, L_0, L_1, L_2, V_0^n, V_1^n, V_2^n, Z^n)$ of the form $p(m, l_0, l_1, l_2, v_0^n, v_1^n, v_2^n z^n | c) = 2^{-n(\tilde{R} + \tilde{R}_1 + \tilde{R}_2)} p(v_0^n, v_1^n, v_2^n | l_0, l_1, l_2, c) \prod_{i=1}^n p_{Z|V_0, V_1, V_2}(z_i | v_{0i}, v_{1i}, v_{2i})$. We again analyze the mutual information between $M$ and $(Z^n, Q^n)$, averaged over codebooks.

$$
\begin{aligned}
I(M; Z^n, Q^n | \mathcal{C}) &= I(M; Z^n | Q^n, \mathcal{C}) \\
&= I(T_1(L_0, L_1), T_2(L_0, L_1), L_0, M; Z^n | Q^n, \mathcal{C}) \\
&\quad - I(T_1(L_0, L_1), T_2(L_0, L_2), L_0; Z^n | M, Q^n, \mathcal{C}) \\
&\le I(V_0^n, V_1^n, V_2^n; Z^n | Q^n, \mathcal{C}) - I(L_0; Z^n | M, Q^n, \mathcal{C}) \\
&\quad - I(T_1(L_0, L_1), T_2(L_0, L_2); Z^n | L_0, Q^n, \mathcal{C}) \\
&\le n I(V_0, V_1, V_2; Z | Q) - H(L_0 | M, Q^n, \mathcal{C}) + H(L_0 | M, Q^n, Z^n, \mathcal{C}) \\
&\quad - I(T_1(L_0, L_1), T_2(L_0, L_2); Z^n | L_0, Q^n, \mathcal{C}) + n\delta(\epsilon).
\end{aligned}
\tag{7}
$$

In the last step, we bound the term $I(V_0^n, V_1^n, V_2^n; Z^n | Q^n, \mathcal{C})$ by the following argument, which is an extension of a similar argument in [1]. For simplicity of notation, let $V = (V_0, V_1, V_2)$. We wish to show that $I(V^n; Z^n | Q^n, \mathcal{C}) \le n I(V; Z) + n\delta(\epsilon)$. Note that $X$ is generated according to $p(x_i | v_i)$. Define $E = 1$ if $(q^n, v^n, z^n)$ are not jointly typical and 0 otherwise, and $N(v) = |\{V_i : V_i = v\}|$. Then,

$$
\begin{aligned}
I(V^n; Z^n | \mathcal{C}, Q^n) &\le 1 + P(E = 0) I(V^\cdot Z^n | \mathcal{C}, E = 0, Q^n) + P(E = 1) I(V^\cdot Z^n | \mathcal{C}, E = 1, Q^n) \\
&\le 1 + P(E = 0) I(V^\cdot Z^n | \mathcal{C}, E = 0, Q^n) \\
&\quad + P(E = 1) n \log |\mathcal{Z}| - P(E = 1) H(Z^n | \mathcal{C}, V^n, Q^n, E = 1) \\
&= 1 + P(E = 0)(H(Z^n | \mathcal{C}, E = 0, Q^n) - H(Z^n | \mathcal{C}, V^n, Q^n) + P(E = 1) n \log |\mathcal{Z}|.
\end{aligned}
$$

Note that $H(Z^n | \mathcal{C}, E = 0, Q^n) \le n H(Z | Q) + n\delta(\epsilon)$. For $H(Z^n | \mathcal{C}, V^n, Q^n, E = 0) = H(Z^n | \mathcal{C}, V^n, E = 0)$, we have

$$
\begin{aligned}
H(Z^n | \mathcal{C}, E = 0, V^n) &\ge \sum_{c, v^n \in \mathcal{T}_\epsilon^{(n)}} P(V^n = v^n, \mathcal{C} = c) H(Z^n | \mathcal{C} = c, V^n = v^n) \\
&= \sum_{c, v^n \in \mathcal{T}_\epsilon^{(n)}} P(V^n = v^n, \mathcal{C} = c) \sum_{i=1}^n H(Z_i | \mathcal{C} = c, V^n = v^n, Z^{i-1}) \\
&\overset{(a)}{=} \sum_{c, v^n \in \mathcal{T}_\epsilon^{(n)}} P(V^n = v^n, \mathcal{C} = c) \sum_{i=1}^n H(Z_i | V_i = v_i) \\
&= \sum_{c, v^n \in \mathcal{T}_\epsilon^{(n)}} P(V^n = v^n, \mathcal{C} = c) \sum_{v \in \mathcal{V}} N(v) H(Z | V = v) \\
&\overset{(b)}{\ge} \sum_{c, v^n \in \mathcal{T}_\epsilon^{(n)}} P(V^n = v^n, \mathcal{C} = c) \left( \sum_{v \in \mathcal{V}} n(p(v) - \delta(\epsilon)) H(Z | V = v) \right) \\
&\ge n H(Z | V) - n\delta'(\epsilon),
\end{aligned}
$$

where $(a)$ follows since given $V_i$, $X_i$ is generated randomly according to $p(x_i | v_i)$ and since the channel is memoryless, $Z_i$ is independent of all other random variables, and $(b)$ follows since $v^n$ is typical, which implies that $N(v) \ge np(v) - n\delta(\epsilon)$. Finally, since the coding scheme satisfies the encoding constraints, the proof is completed by noting that $P(E = 1) \to 0$ as $n \to \infty$ by the law of large numbers and the mutual covering lemma in [11, Chapter 9]).

We now bound each remaining terms in inequality (7) separately. Note that

$$H(L_0|M,Q^n,\mathcal{C}) = n(\tilde{R}-R), \tag{8}$$

$$H(L_0|M,Q^n,Z^n,\mathcal{C}) \overset{(a)}{\leq} n(S_0 - R - I(V_0;Z|Q) + \delta(\epsilon)), \tag{9}$$

where $(a)$ follows by similar steps to the proof of Corollary 1 and application of Lemma 1, which holds if $\mathrm{P}\{(Q^n, V_0^n(L_0), Z^n) \in \mathcal{T}_\epsilon^{(n)}\} \to 1$ as $n \to \infty$ and $S_0 - R \geq I(V_0;Z|Q)$. The first condition follows since $\mathrm{P}\{(Q^n, V_0^n(L_0), V_1^n(L_0, T_1(L_0, L_1)), V_2^n(L_0, T_2(L_0, L_2)), Z^n) \in \mathcal{T}_\epsilon^{(n)}\} \to 1$ as $n \to \infty$. Next, consider

$$
\begin{aligned}
&I(T_1(L_0, L_1), T_2(L_0, L_2); Z^n | L_0, Q^n, \mathcal{C}) \\
&\quad = H(T_1(L_0, L_1), T_2(L_0, L_2) | L_0, Q^n, \mathcal{C}) - H(T_1(L_0, L_1), T_2(L_0, L_2) | L_0, Q^n, Z^n, \mathcal{C}) \\
&\quad \overset{(a)}{=} H(L_1, L_2 | L_0, Q^n, \mathcal{C}) - H(T_1(L_0, L_1), T_2(L_0, L_2) | L_0, Q^n, Z^n, \mathcal{C}) \\
&\quad \geq H(L_1, L_2 | L_0, Q^n, \mathcal{C}) - H(T_1(L_0, L_1) | L_0, Q^n, Z^n, \mathcal{C}) - H(T_2(L_0, L_2) | L_0, Q^n, Z^n, \mathcal{C}), \tag{10}
\end{aligned}
$$

where $(a)$ holds since given the codebook $\mathcal{C}$ and $L_0$, $(T_1, T_2)$ is a one-to-one function of $(L_1, L_2)$. Now,

$$H(L_1, L_2 | L_0, Q^n, \mathcal{C}) = n(\tilde{R}_1 + \tilde{R}_2), \tag{11}$$

$$H(T_1(L_0, L_1) | L_0, Q^n, Z^n, \mathcal{C}) \overset{(b)}{\leq} n(T_1 - I(V_1;Z|V_0) + \delta(\epsilon)), \tag{12}$$

$$H(T_2(L_0, L_2) | L_0, Q^n, Z^n, \mathcal{C}) \overset{(c)}{\leq} n(T_2 - I(V_2;Z|V_0) + \delta(\epsilon)), \tag{13}$$

where $(b)$ and $(c)$ come from the following analysis. First consider

$$
\begin{aligned}
H(T_1(L_0, L_1) | L_0, Q^n, Z^n, \mathcal{C}) &= H(T_1(L_0, L_1) | v_0^n(L_0), Q^n, Z^n, L_0, \mathcal{C}) \\
&\leq H(T_1(L_0, L_1) | V_0^n, Z^n).
\end{aligned}
$$

We now upper bound the term $H(T_1(L_0, L_1) | V_0^n, Z^n)$.

Since $\mathrm{P}\{(Q^n, V_0^n(L_0), V_1^n(L_0, T_1(L_0, L_1)), V_2^n(L_0, T_2(L_0, L_2)), Z^n) \in \mathcal{T}_\epsilon^{(n)}\} \to 1$ as $n \to \infty$, $\mathrm{P}\{(V_0^n(L_0), V_1^n(L_0, T_1(L_0, L_1)), Z^n) \in \mathcal{T}_\epsilon^{(n)}\} \to 1$ as $n \to \infty$. We can therefore apply Lemma 1 to obtain

$$H(T_1(L_0, L_1) | L_0, Q^n, Z^n, \mathcal{C}) \leq n(T_1 - I(V_1;Z|V_0) + \delta(\epsilon)),$$

if $T_1 \geq I(V_1;Z|V_0) + \delta(\epsilon)$.

The term $H(T_2(L_0, L_2)(L_0, L_2) | L_0, Q^n, Z^n, \mathcal{C})$ can be bound using the same steps to give

$$H(T_2(L_0, L_2) | L_0, Q^n, Z^n, \mathcal{C}) \leq n(T_2 - I(V_2;Z|V_0) + \delta(\epsilon)),$$

if $T_2 \geq I(V_2;Z|V_0) + \delta(\epsilon)$.

Substituting from (11), (12), and (13) into (10) yields

$$
\begin{aligned}
&I(T_1(L_0, L_1), T_2(L_0, L_2); Z^n | L_0, Q^n, \mathcal{C}) \\
&\quad \geq n(\tilde{R}_1 + \tilde{R}_2) - n(T_1 - I(V_1;Z|V_0) + \delta(\epsilon)) - n(T_2 - I(V_2;Z|V_0) + \delta(\epsilon)). \tag{14}
\end{aligned}
$$

Substituting inequality (14), together with (8) and (9) into (7) then yields

$$I(M;Z^n|Q^n,\mathcal{C}) \leq n(I(V_1;V_2;Z|V_0) + T_1 + T_2 - \tilde{R}_1 - \tilde{R}_2 - I(V_1;Z|V_0) - I(V_2;Z|V_0) + 3\delta(\epsilon)).$$

Hence, $I(M;Z^n|Q^n,\mathcal{C}) \leq 3n\delta(\epsilon)$ if

$$I(V_1;V_2;Z|V_0) + T_1 + T_2 - \tilde{R}_1 - \tilde{R}_2 - I(V_1;Z|V_0) - I(V_2;Z|V_0) \leq 0.$$

In summary, the rate constraints arising from the analysis of equivocation rate are

$$
\begin{aligned}
S_0 - R &\geq I(V_0;Z|Q), \\
T_1 &\geq I(V_1;Z|V_0), \\
T_2 &\geq I(V_2;Z|V_0), \\
T_1 + T_2 - \tilde{R}_1 - \tilde{R}_2 &\leq I(V_1;Z|V_0) + I(V_2;Z|V_0) - I(V_1;V_2;Z|V_0).
\end{aligned}
$$

Applying Fourier-Motzkin elimination gives

$$R < I(V_0, V_1; Y_1|Q) - I(V_0, V_1; Z|Q),$$
$$R < I(V_0, V_2; Y_2|Q) - I(V_0, V_2; Z|Q),$$
$$2R < I(V_0, V_1; Y_1|Q) + I(V_0, V_2; Y_2|Q) - 2I(V_0; Z|Q) - I(V_1; V_2|V_0)$$

for some $p(q, v_0, v_1, v_2, x) = p(q, v_0)p(v_1, v_2|v_0)p(x|v_1, v_2, v_0)$ such that $I(V_1, V_2; Z|V_0) \leq I(V_1; Z|V_0) + I(V_2; Z|V_0) - I(V_1; V_2|V_0)$.

The proof of Theorem 1 is then completed by observing that the third inequality is redundant. This is seen by summing the first two inequalities to yield

$$2R \leq I(V_0, V_1; Y_1|Q) - I(V_0, V_1; Z|Q) + I(V_0, V_2; Y_2|Q) - I(V_0, V_2; Z|Q)$$
$$= I(V_0, V_1; Y_1|Q) - I(V_0, V_1; Z|Q) + I(V_0, V_2; Y_2|Q) - 2I(V_0; Z|Q) - I(V_1; Z|V_0) - I(V_2; Z|V_0).$$

This inequality is at least as tight as the third inequality because of the constraint on the pmf. This completes the proof of Theorem 1.

*Special Cases:*

We consider several special cases in which the inner bound in Theorem 1 is tight.

*Reversely Degraded Product Broadcast Channel*: As an example of Theorem 1, consider the reversely degraded product broadcast channel with sender $X = (X_1, X_2 \ldots, X_k)$, receivers $Y_j = (Y_{j1}, Y_{j2} \ldots, Y_{jk})$ for $j = 1, 2, 3$, and conditional probability mass functions $p(y_1, y_2, z|x) = \prod_{l=1}^{k} p(y_{1l}, y_{2l}, z_l|x_l)$. In [5], the following lower bound on secrecy capacity is established

$$C_S \geq \min_{j \in \{1,2\}} \sum_{l=1}^{k} [I(U_l; Y_{jl}) - I(U_l; Z_l)]^+. \tag{15}$$

for some $p(u_1, \ldots, u_k, x) = \prod_{l=1}^{k} p(u_l)p(x_l|u_l)$. Furthermore, this lower bound is shown to be optimal when the channel is reversely degraded (with $U_l = X_l$), i.e., when each sub-channel is degraded but not necessarily in the same order. We can show that this result is a special case of Theorem 1. Define the sets of $l$ indexes: $\mathcal{C} = \{l : I(U_l; Y_{1l}) - I(U_l; Z_l) \geq 0, I(U_l; Y_{2l}) - I(U_l; Z_l) \geq 0\}$, $\mathcal{A} = \{l : I(U_l; Y_{1l}) - I(U_l; Z_l) \geq 0\}$ and $\mathcal{B} = \{l : I(U_l; Y_{2l}) - I(U_l; Z_l) \geq 0\}$. Now, setting $V_0 = \{U_l : l \in \mathcal{C}\}$, $V_1 = \{U_l : l \in \mathcal{A}\}$, and $V_2 = \{U_l : l \in \mathcal{B}\}$ in the rate expression of Theorem 1 yields (15). Note that the constraint in Theorem 1 is satisfied for this choice of auxiliary random variables. The expanded equations are as follows:

$$I(V_1, V_2; Z|V_0) = I(U_A, U_B; Z|U_C)$$
$$= I(U_{A\backslash C}, U_{B\backslash C}; Z_{\backslash C})$$
$$= I(U_{A\backslash C}; Z_{,A\backslash C}) + I(U_{B\backslash C}; Z_{,B\backslash C})$$
$$= I(V_1; Z|V_0) + I(V_2; Z|V_0),$$
$$I(V_0, V_1; Y_1) - I(V_0, V_1; Z) = I(U_A; Y_{1,A}) - I(U_A; Z_A),$$
$$I(V_0, V_1; Y_1) - I(V_0, V_1; Z) = I(U_B; Y_{1,A}) - I(U_B; Z_B),$$
$$I(V_1; V_2|V_0) = I(U_{A\backslash C}; U_{B\backslash C}) = 0.$$

*Receivers $Y_1$ and $Y_2$ are less noisy than $Z$*: Recall that in a 2-receiver broadcast channel, a receiver $Y$ is said to be less noisy [14] than a receiver $Z$ if $I(U; Y) \geq I(U; Z)$ for all $p(u, x)$. In this case, we have

$$C_S = \max_{p(x)} \min\{I(X; Y_1) - I(X; Z), I(X; Y_2) - I(X; Z)\}.$$

To show achievability, we set $Q = \emptyset$ and $V_0 = V_1 = V_2 = V_3 = X$ in Theorem 1. The converse follows similar steps to the converse for Proposition 2 in Subsection IV-B given in Appendix IV and we omit it here.

## B. 2-Receivers, 1-Eavesdropper with Common Message

As a generalization of Theorem 1, consider the setting with both common and confidential messages, where we are interested in achieving some equivocation rate for the confidential message rather than asymptotic secrecy. For this setting we can establish the following inner bound on the secrecy capacity region.

*Theorem 2:* An inner bound to the secrecy capacity region of the 2-receiver, 1-eavesdropper broadcast channel with one common and one confidential messages is given by the set of non-negative rate tuples $(R_0, R_1, R_e)$ such that

$$R_0 < I(U; Z),$$
$$R_0 + R_1 < I(U; Z) + \min\{I(V_0, V_1; Y_1|U) - I(V_1; Z|V_0), I(V_0, V_2; Y_2|U) - I(V_2; Z|V_0)\},$$
$$R_0 + R_1 < \min\{I(V_0, V_1; Y_1) - I(V_1; Z|V_0), I(V_0, V_2; Y_2) - I(V_2; Z|V_0)\},$$
$$R_e \leq R_1,$$
$$R_e < \min\{I(V_0, V_1; Y_1|U) - I(V_0, V_1; Z|U), I(V_0, V_2; Y_2|U) - I(V_0, V_2; Z|U)\},$$
$$R_0 + R_e < \min\{I(V_0, V_1; Y_1) - I(V_1, V_0; Z|U), I(V_0, V_2; Y_2) - I(V_2, V_0; Z|U)\},$$
$$R_0 + 2R_e < I(V_0, V_1; Y_1) + I(V_0, V_2; Y_2|U) - I(V_1; V_2|V_0) - 2I(V_0; Z|U),$$
$$R_0 + 2R_e < I(V_0, V_2; Y_2) + I(V_0, V_1; Y_1|U) - I(V_1; V_2|V_0) - 2I(V_0; Z|U),$$
$$R_0 + R_1 + 2R_e < I(V_0, V_2; Y_2|U) - I(V_2; Z|V_0) + I(V_0, V_1; Y_1)$$
$$+ I(V_0, V_2; Y_2|U) - I(V_1; V_2|V_0) - 2I(V_0; Z|U),$$
$$R_0 + R_1 + 2R_e < I(V_0, V_1; Y_1|U) - I(V_1; Z|V_0) + I(V_0, V_2; Y_2)$$
$$+ I(V_0, V_1; Y_1|U) - I(V_1; V_2|V_0) - 2I(V_0; Z|U),$$

for some $p(u, v_0, v_1, v_2, x) = p(u)p(v_0|u)p(v_1, v_2|v_0)p(x|v_0, v_1, v_2)$ such that $I(V_1, V_2; Z|V_0) \leq I(V_1; Z|V_0) + I(V_2; Z|V_0) - I(V_1; V_2|V_0)$.

Note that if we discard the equivocation rate constraints and set $V_0 = V_1 = V_2 = X$, this inner bound reduces to the straightforward extension of the Körner–Marton degraded message set capacity region for the 3 receivers case [8, Corollary 1].

If we take $V_0 = V_1 = V_2 = V$ and $Y_1 = Y_2 = Y$, then we obtain the region consisting of all rate pairs $(R_0, R_1)$ such that

$$R_0 < I(U; Z), \tag{16}$$
$$R_0 + R_1 < I(U; Z) + I(V; Y|U),$$
$$R_0 + R_1 < I(V; Y),$$
$$R_e \leq R_1,$$
$$R_e < I(V; Y|U) - I(V; Z|U),$$
$$R_0 + R_e < I(V; Y) - I(V; Z|U) \tag{17}$$

for some $p(u, v, x) = p(u)p(v|u)p(x|v)$.

This region provides an equivalent characterization of the secrecy capacity region of the 2-receiver broadcast channel with confidential messages [2]. To see this, note that if we tighten the first inequality to $R_0 \leq \min\{I(U; Z), I(U; Y)\}$, the last inequality becomes redundant and the region reduces to the original characterization in [2].

*Proof of Theorem 2*:

The proof of Theorem 2 involves rate splitting for $R_1(= R_1' + R_1'')$. We first establish an inner bound without rate splitting. The proof with rate splitting is given in Appendix III.

*Codebook generation*: Fix $p(u, v_0, v_1, v_2, x)$ and let $R_r \geq 0$ be such that $R_1 - R_e + R_r \geq I(V_0; Z|U) + \delta(\epsilon)$. Randomly and independently generate sequences $u^n(m_0)$, $m_0 \in [1 : 2^{nR_0}]$, each according to $\prod_{i=1}^n p_U(u_i)$. For each $m_0$, randomly and conditionally independently generate sequences $v_0^n(m_0, m_1, m_r)$, $(m_1, m_r) \in [1 : 2^{n(R_1+R_r)}]$, each according to $\prod_{i=1}^n p_{V_0|U}(v_{0i}|u_i)$. For each $(m_0, m_1, m_r)$, generate sequences $v_1^n(m_0, m_1, m_r, t_1)$,

$t_1 \in [1 : 2^{nT_1}]$, each according to $\prod_{i=1}^{n} p_{V_1|V_0}(v_{1i}|v_{0i})$, and partition the set $[1 : 2^{nT_1}]$ into $2^{n\tilde{R}_1}$ equal size bins $\mathcal{B}(m_0, m_1, m_r, l_1)$. Similarly, for each $(m_0, m_1, m_r)$, randomly generate sequences $v_2^n(m_0, m_1, m_r, t_2)$, $t_2 \in [1 : 2^{nT_2}]$ each according to $\prod_{i=1}^{n} p_{V_2|V_0}(v_{2i}|v_{0i})$ and partition the set $[1 : 2^{nT_2}]$ into $2^{n\tilde{R}_2}$ bins $\mathcal{B}(m_0, m_1, m_r, l_2)$. Finally, for each product bin $\mathcal{B}(l_1) \times \mathcal{B}(l_2)$, find a jointly typical sequence pair $(v_1^n(m_0, m_1, m_r, t_1(l_1)), v_2^n(m_0, m_1, m_r, t_2(l_2)))$. If there is more than 1 pair, we randomly and uniformly pick a pair from the set of jointly typical pairs. As before, the probability of error approaches zero as $n \to \infty$ if

$$\tilde{R}_1 + \tilde{R}_2 < T_1 + T_2 - I(V_1; V_2|V_0) - \delta(\epsilon).$$

*Encoding*: To send a message pair $(m_0, m_1)$, the encoder first chooses a random index $m_r \in [1 : 2^{nR_r}]$ and then the sequence pair $(u^n(m_0), v_0^n(m_1, m_r, m_0))$. It then randomly chooses a product bin indices $(L_1, L_2)$ and selects the jointly typical sequence pair
$(v_1^n(m_0, m_1, m_r, t_1(L_1)), v_2^n(m_0, m_1, m_r, t_2(L_2))$ in it. Finally, it generates a codeword $X^n$ at random according to $\prod_{i=1}^{n} p_{X|V_0, V_1.V_2}(x_i|v_{0i}, v_{1i}, v_{2i})$.

*Decoding and analysis of the probability of error*: Receiver $Y_1$ finds $(m_0, m_1)$ indirectly by looking for the unique $(m_0, \hat{l}_0)$ such that $(u^n(m_0), v_0^n(m_0, \hat{l}_0), v_1^n(m_0, \hat{l}_0, l_1)) \in \mathcal{T}_\epsilon^{(n)}$ for some $l_1 \in [1 : 2^{nT_1}]$. Similarly, receiver $Y_2$ finds $(m_0, m_1)$ indirectly by looking for the unique $(m_0, \hat{l}_0)$ such that $(u^n(m_0), v_0^n(m_0, \hat{l}_0), v_1^n(m_0, \hat{l}_0, l_2)) \in \mathcal{T}_\epsilon^{(n)}$ for some $l_2 \in [1 : 2^{nT_2}]$. Receiver $Z$ finds $m_0$ directly by decoding $U$. These steps succeed with probability of error approaching zero as $n \to \infty$ if

$$R_0 + R_1 + T_1 + R_r < I(V_0, V_1; Y_1) - \delta(\epsilon),$$
$$R_1 + T_1 + R_r < I(V_0, V_1; Y_1|U) - \delta(\epsilon),$$
$$R_0 + R_1 + T_2 + R_r < I(V_0, V_1; Y_2) - \delta(\epsilon),$$
$$R_1 + T_2 + R_r < I(V_0, V_1; Y_2|U) - \delta(\epsilon),$$
$$R_0 < I(U; Z).$$

*Analysis of equivocation rate*: We consider the equivocation rate averaged over codes. We will show that a part of the message $M_{1p}$ can be kept asymptotically secret from the eavesdropper as long as rate constraints on $R_e$ and $R_1$ are satisfied. Let $R_1 = R_{1p} + R_{1c}$ and $R_e = R_{1p}$.

$$\begin{aligned} H(M_1|Z^n, \mathcal{C}) &\geq H(M_{1p}|Z^n, M_0, \mathcal{C}) \\ &= H(M_{1p}) - I(M_{1p}; Z^n|M_0, \mathcal{C}) \quad\quad\quad (18) \\ &\stackrel{(a)}{\geq} H(M_{1p}) - 3n\delta(\epsilon) \\ &= n(R_1 - I(V_0; Z|U)) - 3n\delta(\epsilon). \end{aligned}$$

This implies that $R_e \leq R_1 - I(V_0; Z|U) - 3\delta(\epsilon)$ is achievable.

To prove step $(a)$, consider

$$\begin{aligned} I(M_{1p}; Z^n|M_0, \mathcal{C}) &= I(T_1(L_1), T_2(L_2), M_{1p}, M_{1c}, M_r; Z^n|M_0, \mathcal{C}) - I(T_1(L_1), T_2(L_2), M_{1c}, M_r; Z^n|M_{1p}, M_0, \mathcal{C}) \\ &\stackrel{(b)}{\leq} I(V_0^n, V_1^n, V_2^n; Z^n|M_0, \mathcal{C}) - I(M_{1c}, M_r; Z^n|M_{1p}, M_0, \mathcal{C}) - I(T_1(L_1), T_2(L_2); Z^n|M_1, M_0, M_r, \mathcal{C}) \\ &\stackrel{(c)}{\leq} I(V_0^n, V_1^n, V_2^n; Z^n|U^n, \mathcal{C}) - I(M_{1c}, M_r; Z^n|M_{1p}, M_0, \mathcal{C}) - I(T_1(L_1), T_2(L_2); Z^n|M_1, M_0, M_r, \mathcal{C}) \\ &\leq nI(V_0, V_1, V_2; Z|U) + n\delta(\epsilon) - H(M_{1c}, M_r|M_{1p}, U^n, \mathcal{C}) + H(M_{1c}, M_r|M_{1p}, M_0, Z^n, \mathcal{C}) \\ &\quad - I(T_1(L_1), T_2(L_2); Z^n|M_1, M_0, M_r, \mathcal{C}) \\ &\leq nI(V_0, V_1, V_2; Z|U) + n\delta(\epsilon) - n(R_1 - R_e + R_r) + H(M_{1c}, M_r|M_{1p}, M_0, Z^n, \mathcal{C}) \\ &\quad - I(T_1(L_1), T_2(L_2); Z^n|M_1, M_0, M_r, \mathcal{C}), \end{aligned}$$

where $(b)$ follows by the data processing inequality and $(c)$ follows by the observation that $U^n$ is a function of $(\mathcal{C}, M_0)$ and $(\mathcal{C}, M_0) \to (\mathcal{C}, U^n, V^n) \to Z^n$. Following the analysis of the equivocation rate terms in Theorem 1 and using Lemma 1, the remaining terms can be bounded by

$$H(M_{1c}, M_r|M_{1p}, M_0, Z^n, \mathcal{C}) \leq H(M_{1c}, M_r|M_{1p}, U^n, Z^n)$$

$$\leq n(R_1 - R_e + R_r) - nI(V_0; Z|U) + n\delta(\epsilon),$$

$$I(T_1(L_1), T_2(L_2); Z^n | M_1, M_0, M_r, \mathcal{C}) = H(T_1(L_1), T_2(L_2) | M_1, M_0, M_r, \mathcal{C}) - H(T_1(L_1), T_2(L_2) | M_1, M_0, M_r, \mathcal{C}, Z^n)$$

$$= n(\tilde{R}_1 + \tilde{R}_2) - H(T_1(L_1), T_2(L_2) | M_1, M_0, M_r, \mathcal{C}, Z^n)$$

$$\overset{(a)}{=} n(\tilde{R}_1 + \tilde{R}_2) - H(T_1(L_1), T_2(L_2) | M_r, M_1, M_0, V_0^n, \mathcal{C}, Z^n)$$

$$\geq n(\tilde{R}_1 + \tilde{R}_2) - H(T_1(L_1), T_2(L_2) | V_0^n, Z^n)$$

$$\geq n(\tilde{R}_1 + \tilde{R}_2 - T_1 - T_2) + n(I(V_1; Z|V_0) + I(V_2; Z|V_0)) - 2n\delta(\epsilon),$$

if $T_1 \geq I(V_1; Z|V_0) + \delta(\epsilon)$, and $T_2 \geq I(V_2; Z|V_0) + \delta(\epsilon)$. Step $(a)$ follows from the observation that $V_0^n$ is a function of $(\mathcal{C}, M_0, M_1)$.

Thus, we have

$$I(M_{1p}; Z^n | M_0, \mathcal{C}) \leq I(V_1, V_2; Z|V_0) - I(V_1; Z|V_0) - I(V_2; Z|V_0) + n(T_1 + T_2 - \tilde{R}_1 - \tilde{R}_2) + 4n\delta(\epsilon).$$

Hence, $I(M_{1p}; Z^n | M_0, \mathcal{C}) \leq 4n\delta(\epsilon)$ if

$$I(V_1; V_2; Z|V_0) + T_1 + T_2 - \tilde{R}_1 - \tilde{R}_2 - I(V_1; Z|V_0) - I(V_2; Z|V_0)) \leq 0.$$

Substituting back into (18) shows that

$$H(M_1 | Z^n, \mathcal{C}) \geq n(R_1 - I(V_0; Z|U) - 4n\delta(\epsilon).$$

The equivocation rate constraints on the rates are

$$R_e \leq R_1,$$
$$R_r \geq 0,$$
$$R_1 - R_e + R_r \geq I(V_0; Z|U),$$
$$T_1 \geq I(V_1; Z|V_0),$$
$$T_2 \geq I(V_2; Z|V_0).$$

Using Fourier-Motzkin elimination then gives us an inner bound for the case without rate splitting. The proof with rate splitting on $R_1$ is given in Appendix III.


*Special Case:*

We show that the inner bound in Theorem 2 is tight when both $Y_1$ and $Y_2$ are less noisy than $Z$.

*Proposition 2:* When both $Y_1$ and $Y_2$ are less noisy than $Z$, the 2-receiver, 1-eavesdropper secrecy capacity region is given by the set of $(R_0, R_1, R_e)$ tuples such that

$$R_0 \leq I(U; Z),$$
$$R_1 \leq \min\{I(X; Y_1|U), I(X; Y_2|U)\},$$
$$R_e \leq [\min\{R_1, I(X; Y_1|U) - I(X; Z|U), I(X; Y_2|U) - I(X; Z|U)\}]^+$$

for some $p(u, x)$.

Achievability follows by setting $V_0 = V_1 = V_2 = X$ in Theorem 2 and using the fact that $Y_1$ and $Y_2$ are less noisy than $Z$, which allows us to assume without loss of generality that $R_0 \leq \min\{I(U; Z), I(U; Y_1), I(U; Y_2)\}$. The set of inequalities then reduce to

$$R_0 < I(U; Z),$$
$$R_0 + R_1 < I(U; Z) + \min\{I(X; Y_1|U), I(X; Y_2|U)\},$$
$$R_e \leq R_1,$$
$$R_e < \min\{I(X; Y_1|U) - I(X; Z|U), I(X; Y_2|U) - I(X; Z|U)\}.$$

Since the region in Proposition 2 is a subset of the above region, we have established the achievability part of the proof. Achievability in this case, however, is a straightforward extension of Csiszár and Körner and does not require Marton coding. For the converse, we use the identification $U_i = (M_0, Z^{i-1})$. With this identification, the $R_0$ inequality follows trivially. The $R_1$ and $R_e$ inequalities follow from standard methods and a technique in [8, Proposition 11]. The details are given in Appendix IV.

## V. 1-RECEIVER, 2-EAVESDROPPERS WIRETAP CHANNEL

We now consider the case where the confidential message $M_1$ is to be sent only to $Y_1$ and kept hidden from the eavesdroppers $Z_2$ and $Z_3$. All three receivers $Y_1, Z_2, Z_3$ require a common message $M_0$. For simplicity, we only consider the special case of multilevel broadcast channel [10], where $p(y_1, z_2, z_3|x) = p(y_1, z_3|x)p(z_2|y_1)$. In [8], it was shown that the capacity region (without secrecy) is the set of rate pairs $(R_0, R_1)$ such that

$$R_0 < \min\{I(U; Z_2), I(U_3; Z_3)\},$$
$$R_1 < I(X; Y_1|U),$$
$$R_0 + R_1 < I(U_3; Z_3) + I(X; Y_1|U_3)$$

for some $p(u)p(u_3|u)p(x|u_3)$. We extend this result to obtain inner and outer bounds on the secrecy capacity region.

*Proposition 3:* An inner bound to the secrecy capacity region of the 1-receiver, 2-eavesdropper multilevel broadcast channel with common and confidential messages is is given by the set of rate tuples $(R_0, R_1, R_{e2}, R_{e3})$ such that

$$R_0 < \min\{I(U; Z_2), I(U_3; Z_3)\},$$
$$R_1 < I(V; Y_1|U),$$
$$R_0 + R_1 < I(U_3; Z_3) + I(V; Y_1|U_3),$$
$$R_{e2} \le \min\{R_1, I(V; Y_1|U) - I(V; Z_2|U)\},$$
$$R_{e2} \le [I(U_3; Z_3) - R_0 - I(U_3; Z_2|U)]^+ + I(V; Y_1|U_3) - I(V; Z_2|U_3),$$
$$R_{e3} \le \min\{R_1, [I(V; Y_1|U_3) - I(V; Z_3|U_3)]^+\},$$
$$R_{e2} + R_{e3} \le R_1 + I(V; Y_1|U_3) - I(V; Z_2|U_3),$$

for some $p(u, u_3, v, x) = p(u)p(u_3|u)p(v|u_3)p(x|v)$.

It can be shown that setting $Y_1 = Z_2 = Y$ and $Z_3 = Z$ gives an alternative characterization of the secrecy capacity of the broadcast channel with confidential messages.

*Proof of achievability*: We break down the proof of Proposition 3 into four cases and give the analysis of the first case in detail. The analyses for the rest of the cases are similar and we therefore we only provide a sketch in Appendix V. Furthermore, in all cases, we assume that $R_1 \ge \min\{I(V; Y_1|U_3) - I(V; Z_2|U_3), [I(V; Y_1|U_3) - I(V; Z_3|U_3)]^+\}$. It is easy to see from our proof that if this inequality does not hold, then we achieve equivocation rates of $R_{e2} = R_{e3} = R_1$ for any rate pair $(R_0, R_1)$ satisfying the inequalities in the proposition. The four cases are:

- Case 1: $I(U_3; Z_3) - R_0 - I(U_3; Z_2|U) \ge 0$, $I(V; Y_1|U_3) - I(V; Z_2|U_3) \le I(V; Y_1|U_3) - I(V; Z_3|U_3)$ and $R_{e3} \ge I(V; Y_1|U_3) - I(V; Z_2|U_3)$;

- Case 2: $I(U_3; Z_3) - R_0 - I(U_3; Z_2|U) \ge 0$, $I(V; Y_1|U_3) - I(V; Z_2|U_3) \le I(V; Y_1|U_3) - I(V; Z_3|U_3)$ and $R_{e3} \le I(V; Y_1|U_3) - I(V; Z_2|U_3)$;

- Case 3: $I(U_3; Z_3) - R_0 - I(U_3; Z_2|U) \ge 0$, $I(V; Y_1|U_3) - I(V; Z_2|U_3) \ge I(V; Y_1|U_3) - I(V; Z_3|U_3)$. In this case, since we consider only the case of $R_1 \ge I(V; Y_1|U_3) - I(V; Z_3|U_3)$, we will see that an equivocation rate of $R_{e3} = I(V; Y_1|U_3) - I(V; Z_3|U_3)$ can be achieved;

- Case 4: $I(U_3; Z_3) - R_0 - I(U_3; Z_2|U) \le 0$.

Now, consider Case 1, where $I(U_3; Z_3) - R_0 - I(U_3; Z_2|U) \ge 0$, $I(V; Y_1|U_3) - I(V; Z_2|U_3) \le I(V; Y_1|U_3) - I(V; Z_3|U_3)$ and $R_{e3} \ge I(V; Y_1|U_3) - I(V; Z_2|U_3)$.

14

*Codebook generation*: Fix $p(u, u_3, v, x) = p(u)p(u_3|u)p(v|u_3)p(x|v)$. Let $R_1 = R_{10}^o + R_{10}^s + R_{11}' + R_{11}'' + R_{11}^o$. Let $R_0^r \geq 0$ and $R_1^r \geq 0$ be the randomization rates introduced by the encoder. These are not part of the message rate. Let $\tilde{R}_{10} = R_{10}^o + R_{10}^s + R_0^r$ and $\tilde{R}_{11} = R_{11}' + R_{11}'' + R_{11}^o + R_1^r$.

Randomly and independently generate sequences $u^n(m_0)$, $m_0 \in [1 : 2^{nR_0}]$, each according to $\prod_{i=1}^n p_U(u_i)$. For each $m_0$, randomly and conditionally independently generate sequences $u_3^n(m_0, l_0)$, $l_0 \in [1 : 2^{n\tilde{R}_{10}}]$, each according to $\prod_{i=1}^n p_{U_3|U}(u_{3i}|u_i)$. For each $(m_0, l_0)$, randomly and conditionally independently generate sequences $v^n(m_0, l_0, l_1)$, $l_1 \in [1 : 2^{n\tilde{R}_{11}}]$, each according to $\prod_{i=1}^n p_{V|U_3}(v_i|u_{3i})$.

*Encoding*: To send a message $(m_0, m_1)$, we split $m_1$ into sub-messages with the corresponding rates given in the codebook generation step and generate the randomization messages $(m_{10}^r, m_{11}^r)$ uniformly at random from the set $[1 : 2^{nR_0^r}] \times [1 : 2^{nR_1^r}]$. We then select the sequence $v^n(m_0, l_0, l_1)$ corresponding to $(m_0, m_1, m_{10}^r, m_{11}^r)$ and send $X^n$ generated according to $\prod_{i=1}^n p_{X|V}(x_i|v_i(l_1, l_0, m_0))$.

*Decoding and analysis of the probability of error*: Receiver $Y_1$ finds $(m_0, m_1)$ by decoding $(U, U_3, V)$, $Z_2$ finds $m_0$ by decoding $U$, and $Z_3$ finds $m_0$ indirectly through $(U, U_3)$. The probability of error goes to zero as $n \to \infty$ if

$$R_0 \leq I(U; Z_2),$$
$$R_0 + R_{10}^o + R_0^r + R_{10}^s < I(U_3; Z_3) - \delta(\epsilon),$$
$$R_{10}^s + R_{10}^o + R_0^r < I(U_3; Y_1|U) - \delta(\epsilon),$$
$$R_{11}' + R_{11}'' + R_{11}^o + R_1^r < I(V; Y_1|U_3) - \delta(\epsilon).$$

*Analysis of equivocation rates*: We show that the following equivocation rates are achievable.

$$R_{e2} = R_{10}^s + R_{11}' - \delta(\epsilon),$$
$$R_{e3} = R_{11}' + R_{11}'' - \delta(\epsilon).$$

It is straightforward to show that the stated equivocation rate $R_{e3}$ is achievable if

$$R_1^r + R_{11}^o > I(V; Z_3|U_3) + \delta(\epsilon).$$

The analysis of the $H(M_1|Z_2^n, \mathcal{C})$ term is slightly more involved. Consider

$$I(M_{10}^s, M_{11}'; Z_2^n|M_0, \mathcal{C}) = I(L_0, L_1; Z_2^n|\mathcal{C}, M_0) - I(L_0, L_1; Z_2^n|\mathcal{C}, M_0, M_{10}^s, M_{11}')$$
$$\leq I(V^n; Z_2^n|\mathcal{C}, U^n) - I(L_0; Z_2^n|\mathcal{C}, M_0, M_{10}^s, M_{11}') - I(L_1; Z_2^n|\mathcal{C}, M_0, L_0, M_{11}')$$
$$\leq \sum_{i=1}^n I(V; Z_2|U) - I(L_0; Z_2^n|\mathcal{C}, M_0, M_{10}^s, M_{11}') - I(L_1; Z_2^n|\mathcal{C}, M_0, L_0, M_{11}').$$

Now consider the second and third terms. We have

$$I(L_0; Z_2^n|\mathcal{C}, M_0, M_{10}^s, M_{11}') = H(L_0|\mathcal{C}, M_0, M_{10}^s, M_{11}') - H(L_0|\mathcal{C}, M_0, M_{10}^s, M_{11}', Z_2^n, U^n)$$
$$\geq n(\tilde{R}_{10} - R_{10}^s) - H(L_0|\mathcal{C}, M_{10}^s, Z_2^n, U^n)$$
$$\geq n(I(U_3; Z_2|U) - \delta(\epsilon)).$$

The last step follows from Lemma 1, which holds if

$$\tilde{R}_{10} - R_{10}^s = R_{10}^o + R_0^r$$
$$\geq I(U_3; Z_2|U) + \delta(\epsilon).$$

For the third term, we have

$$I(L_1; Z_2^n|\mathcal{C}, M_0, L_0, M_{11}') = H(L_1|\mathcal{C}, M_0, L_0, M_{11}') - H(L_1|\mathcal{C}, M_0, L_0, M_{11}', Z_2^n, U^n)$$
$$\geq n(\tilde{R}_{11} - R_{11}') - H(L_1|\mathcal{C}, M_{11}', Z_2^n, U^n)$$
$$\geq n(\tilde{R}_{11} - R_{11}') - n(\tilde{R}_{11} - R_{11}' - I(V; Z_2|U_3) + \delta(\epsilon)).$$

In the last step, we again apply Lemma 1, which holds if

$$R''_{11} + R^o_{11} + R^r_1 \geq I(V; Z_|U_3) + \delta(\epsilon).$$

In summary, the inequalities for Case 1 are as follows:

*Decoding Constraints*: (with $R_0 \leq I(U; Z_2)$ omitted since this inequality appears in the final rate-equivocation region and does not contain the auxiliary rates to be eliminated.)

$$R_0 + R^o_{10} + R^r_0 + R^s_{10} < I(U_3; Z_3),$$
$$R^s_{10} + R^o_{10} + R^r_0 < I(U_3; Y_1|U),$$
$$R'_{11} + R''_{11} + R^o_{11} + R^r_1 < I(V; Y_1|U_3).$$

*Equivocation rate constraints*:

$$R^o_{10} + R^r_o > I(U_3; Z_2|U),$$
$$R''_{11} + R^r_1 + R^o_{11} > I(V; Z_2|U_3),$$
$$R^r_1 + R^o_{11} > I(V; Z_3|U_3).$$

*Greater than or equal to zero constraints*:

$$R^o_{10}, R^o_0, R'_{11}, R''_{11}, R^n_1, R^r_0 \geq 0.$$

*Equality constraints*:

$$R_1 = R^o_{10} + R^s_{10} + R'_{11} + R''_{11} + R^o_{11},$$
$$R_{e2} = R^s_{10} + R'_{11},$$
$$R_{e3} = R'_{11} + R''_{11}.$$

Applying Fourier-Motzkin elimination yields the rate-equivocation region for Case one. Sketch of achievability for the other cases are given in Appendix V.

We now establish an outer bound and use it to show that the inner bound in Proposition 3 is tight in several special cases. In contrast to the case with no secrecy constraint [8], the assumption of a stochastic encoder makes it difficult to match our inner and outer bounds in general.

*Proposition 4:* An outer bound on the secrecy capacity of the multilevel 3-receiver broadcast channel with one common and one confidential messages is given by the set of rate tuples $(R_0, R_1, R_{e2}, R_{e3})$ such that

$$R_0 \leq \min\{I(U; Z_2), I(U_3; Z_3)\},$$
$$R_1 \leq I(V; Y_1|U),$$
$$R_0 + R_1 \leq I(U_3; Z_3) + I(V; Y_1|U_3),$$
$$R_{e2} \leq I(X; Y_1|U) - I(X; Z_2|U),$$
$$R_{e2} \leq [I(U_3; Z_3) - R_0 - I(U_3; Z_2|U)]^+ + I(X; Y_1|U_3) - I(X; Z_2|U_3),$$
$$R_{e3} \leq [I(V; Y_1|U_3) - I(V; Z_3|U_3)]^+$$

for some $p(u, u_3, v, x) = p(u)p(u_3|u)p(v|u_3)p(x|v)$.
Proof of this Proposition uses a combination of standard converse techniques from [15], [16], and [2] and given in Appendix VI.

*Remark 5.1*: As we can see in the inequalities governing $R_{e2}$ in both the inner and outer bounds, there is a tradeoff between the common message rate and the equivocation rate at receiver $Z_2$. A higher common message rate limits the number of codewords that can be generated to confuse the eavesdropper.

*Special Cases*

Using Propositions 3 and 4, we can establish the secrecy capacity region for the following special cases.
$Y_1$ *more capable than* $Z_3$ *and* $Z_3$ *more capable than* $Z_2$: If $Y_1$ is *more capable* [15] than $Z_3$ and $Z_3$ is more capable than $Z_2$, the capacity region is given by:

$$R_0 \leq \min\{I(U;Y_2), I(U_3;Z_3)\},$$
$$R_1 \leq I(X;Y_1|U),$$
$$R_0 + R_1 \leq I(U_3;Z_3) + I(X;Y_1|U_3),$$
$$R_{e2} \leq I(X;Y_1|U) - I(X;Z_2|U),$$
$$R_{e2} \leq [I(U_3;Z_3) - R_0 - I(U_3;Z_2|U)]^+ + I(X;Y_1|U_3) - I(X;Z_2|U_3),$$
$$R_{e3} \leq I(X;Y_1|U_3) - I(X;Z_3|U_3)$$

for some $p(u, u_3, x) = p(u)p(u_3|u)p(x|u_3)$.

Achievability follows directly from setting $V = X$ and observing that since $Z_3$ is more capable than $Z_2$, the inequality $R_{e2} + R_{e3} \leq R_1 + I(X;Y_1|U_3) - I(X;Z_2|U_3)$ is redundant since $I(X;Y_1|U_3) - I(X;Z_2|U_3) \geq I(X;Y_1|U_3) - I(X;Z_3|U_3)$ from the more capable condition. For the converse, observe that since $Y_1$ is more capable than $Z_3$, we have

$$I(V;Y_1|U_3) - I(V;Z_3|U_3) = I(V,X;Y_1|U_3) - I(V,X;Z_3|U_3) - I(X;Y_1|V) + I(X;Z_3|V)$$
$$\leq I(X;Y_1|U_3) - I(X;Z_3|U_3).$$

*One eavesdropper*: Here, we consider the two scenarios where either $Z_2$ or $Z_3$ is an eavesdropper and the other receiver is neutral, i.e., there is no constraint on its equivocation rate, but it still decodes a common message. The secrecy capacity regions for these two scenarios are as follows.

$Z_3$ *is neutral*: The secrecy capacity region is the set of rate tuples $(R_0, R_1, R_{e2})$ such that

$$R_0 \leq \min\{I(U;Y_2), I(U_3;Z_3)\},$$
$$R_1 \leq I(X;Y_1|U),$$
$$R_0 + R_1 \leq I(U_3;Z_3) + I(X;Y_1|U_3),$$
$$R_{e2} \leq I(X;Y_1|U) - I(X;Z_2|U),$$
$$R_{e2} \leq [I(U_3;Z_3) - R_0 - I(U_3;Z_2|U)]^+ + I(X;Y_1|U_3) - I(X;Z_2|U_3)$$

for some $p(u, u_3, x) = p(u)p(u_3|u)p(x|u_3)$.

$Z_2$ *is neutral*: The secrecy capacity region is the set of rate tuples $(R_0, R_1, R_{e3})$ such that

$$R_0 \leq \min\{I(U;Z_2), I(U_3;Z_3)\},$$
$$R_1 \leq I(V;Y_1|U),$$
$$R_0 + R_1 \leq I(U_3;Z_3) + I(V;Y_1|U_3),$$
$$R_{e3} \leq [I(V;Y_1|U_3) - I(V;Z_3|U_3)]^+$$

for some $p(u, u_3, v, x) = p(u)p(u_3|u)p(v|u_3)p(x|v)$.

## VI. Conclusion

We presented inner and outer bounds on the secrecy capacity region of the 3-receiver broadcast channel with common and confidential messages that are strictly larger than straightforward extensions of the Csiszár–Körner 2-receiver region. We considered the 2-receiver, 1-eavesdropper and the 1-receiver, 2-eavesdroppers cases. For the first case, we showed that additional superposition encoding, whereby a codeword is picked at random from a pre-generated codebook can increase the achievable rate by allowing the legitimate receiver to indirectly decode the message without sacrificing secrecy. A general lower bound on the secrecy capacity is then obtained by combining superposition encoding and indirect decoding with Marton coding. This lower bound is shown to be tight for the

reversely degraded product channel and when both $Y_1$ and $Y_2$ are less noisy than the eavesdropper. The lower bound was generalized in Theorem 2 to obtain an inner bound on the secrecy capacity region for the 2-receiver, 1 eavesdropper case. For the case where both $Y_1$ and $Y_2$ are less noisy than the eavesdropper, we again show that our inner bound gives the secrecy capacity region.

We then established inner and outer bounds on the secrecy capacity region for the 1-receiver, 2-eavesdroppers multilevel wiretap channel. The inner bound and outer bounds are shown to be tight for several special cases. In the results for both setups, we observe a tradeoff between the common message rate and the eavesdropper equivocation rates. A higher common message rate limits the number of codewords that can be generated to confuse the eavesdroppers about the confidential message. In addition, in the second setup, a higher common message rate can potentially reduce the equivocation rate of one eavesdropper while leaving the equivocation rate at the other eavesdropper unchanged.

REFERENCES

[1] A. D. Wyner, "The wire-tap channel," *Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, 1975.
[2] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Info. Theory*, vol. IT-24, pp. 339–348, May 1978.
[3] R. Liu, I. Maric, P. Spasojevic, and R. D. Yates, "Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions," *IEEE Trans. Info. Theory*, vol. 54, no. 6, pp. 2493–2507, June 2008.
[4] G. Bagherikaram, A. S. Motahari, and A. K. Khandani, "The secrecy rate region of the broadcast channel," in *46th Annual Allerton Conference on Communication, Control and Computing*, Sept 2008.
[5] A. Khisti, A. Tchamkerten, and G. Wornell, "Secure broadcasting over fading channels," *IEEE Trans. Info. Theory*, vol. 54, no. 6, pp. 2453–2469, June 2008.
[6] E. Ekrem and S. Ulukus, "Secrecy capacity of a class of broadcast channels with an eavesdropper," *EURASIP Journal on Wireless Communications and Networking*, Oct 2009.
[7] Y. Liang, H. V. Poor, and S. Shamai, "Information theoretic security," *Foundations and Trends in Communications and Information Theory*, vol. 5, no. 4–5, pp. 355–580, 2008.
[8] C. Nair and A. E. Gamal, "The capacity region of a class of 3-receiver broadcast channels with degraded message sets," *IEEE Trans. Info. Theory*, 2008, submitted. Available online at http://arxiv.org/abs/0712.3327.
[9] Y. K. Chia and A. E. Gamal, "3-receiver broadcast channels with common and confidential messages," in *Proc. IEEE International Symposium on Information Theory*, Seoul, South Korea, June/July 2009.
[10] S. Borade, L. Zheng, and M. Trott, "Multilevel broadcast networks," in *International Symposium on Information Theory*, 2007.
[11] A. El Gamal and Y. H. Kim, "Lectures on network information theory," 2010, available online at ArXiv: http://arxiv.org/abs/1001.3404.
[12] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Info. Theory*, vol. 25, no. 3, pp. 306–311, May 1979.
[13] A. E. Gamal and E. C. van der Meulen, "A proof of marton's coding theorem for the discrete memoryless broadcast channel," *IEEE Transactions on Information Theory*, vol. 27, no. 1, pp. 120–121, 1981.
[14] J. Körner and K. Marton, "Comparison of two noisy channels," in *Topics in Information Theory (Second Colloq., Keszthely, 1975)*, 1977, pp. 411–423.
[15] A. E. Gamal, "The capacity of a class of broadcast channels," *IEEE Trans. Info. Theory*, vol. 25, no. 2, pp. 166–169, Mar 1979.
[16] ——, "The feedback capacity of degraded broadcast channels (corresp.)," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 379–381, May 1978.
[17] T. Cover and J. Thomas, *Elements of Information Theory 2nd edition*. Wiley Interscience, July 2006.

APPENDIX I
PROOF OF LEMMA 1

First, define $N(U^n, Z^n) = |\{k \in [1 : 2^{nS}] : (U^n, V^n(k), Z^n) \in \mathcal{T}_\epsilon^{(n)}\}|$. Next, we define the following "error" events. Let $E_1(U^n, Z^n) = 1$ if $\{N(U^n, Z^n) \geq (1 + \delta_1(\epsilon))2^{n(S-I(V;Z|U)+\delta(\epsilon))}\}$ and $E_1 = 0$ otherwise. Let $E = 0$ if $(U^n, V^n(L), Z^n) \in \mathcal{T}_\epsilon^{(n)}$ and $E_1(U^n, Z^n, L) = 0$, and $E = 1$ otherwise. We now show that if $S \geq I(V; Z|U) + \delta(\epsilon)$, then $\mathrm{P}\{E = 1\} \to 0$ as $n \to \infty$. By the union of events bound,

$$\mathrm{P}\{E = 1\} \leq \mathrm{P}\{(U^n, V^n(L), Z^n) \notin \mathcal{T}_\epsilon^{(n)}\} + \mathrm{P}\{E_1(U^n, Z^n, L) = 1\}.$$

The first term tends to zero as $n \to \infty$ by assumption. The second term is bounded as follows

$$
\begin{aligned}
\mathrm{P}\{E_1(U^n, Z^n) = 1\} &= \sum_{u^n \in \mathcal{T}_\epsilon^{(n)}} p(u^n) \, \mathrm{P}\{(E_1(U^n, Z^n) = 1)|U^n = u^n\} \\
&= \sum_{u^n \in \mathcal{T}_\epsilon^{(n)}} \sum_{z^n \in \mathcal{T}_\epsilon^{(n)}(Z|U)} p(u^n) \, \mathrm{P}\{(E_1(u^n, Z^n) = 1) \cap (Z^n = z^n)|U^n = u^n\} \\
&= \sum_{u^n \in \mathcal{T}_\epsilon^{(n)}} \sum_{z^n \in \mathcal{T}_\epsilon^{(n)}(Z|U)} p(u^n) \, \mathrm{P}\{(E_1(u^n, z^n) = 1) \cap (Z^n = z^n)|U^n = u^n\} \\
&\leq \sum_{u^n \in \mathcal{T}_\epsilon^{(n)}} p(u^n) \sum_{z^n \in \mathcal{T}_\epsilon^{(n)}(Z|U)} \mathrm{P}\{(E_1(u^n, z^n) = 1)|U^n = u^n\}.
\end{aligned}
$$

Now, $\mathrm{P}\{E_1(u^n, z^n) = 1|U^n = u^n\} = \mathrm{P}\{N(u^n, z^n) \geq (1 + \delta_1(\epsilon))2^{n(S-I(V;Z|U)+\delta(\epsilon))}|U^n = u^n\}$. Define $X_k = 1$ if $(u^n, V^n(k), z^n) \in \mathcal{T}_\epsilon^{(n)}$ and 0, otherwise. We note that $X_k$, $k \in [1 : 2^{nS}]$, are i.i.d. Bernoulli $p$ random variables, where $2^{-n(I(V;Z|U)+\delta(\epsilon))} \leq p \leq 2^{-n(I(V;Z|U)-\delta(\epsilon))}$. We have

$$
\begin{aligned}
&\mathrm{P}\{N(u^n, z^n) \geq (1 + \delta_1(\epsilon))2^{n(S-I(V;Z|U)+\delta(\epsilon))}|U^n = u^n\} \\
&\leq \mathrm{P}\left\{ \sum_{k=1}^{2^{nS}} X_k \geq (1 + \delta_1(\epsilon))2^{nS}p|U^n = u^n \right\}.
\end{aligned}
$$

Applying the Chernoff Bound (e.g., see [11, Appendix B]), we have

$$
\begin{aligned}
\mathrm{P}\left\{ \sum_{k=1}^{2^{nS}} X_k \geq (1 + \delta_1(\epsilon))2^{nS}p|U^n = u^n \right\} &\leq \exp(-2^{nS}p\delta_1^2(\epsilon)/4) \\
&\leq \exp(-2^{n(S-I(V;Z|U)-\delta(\epsilon))}\delta_1^2(\epsilon)/4).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\mathrm{P}\{E_1(U^n, Z^n) = 1\} &\leq \sum_{u^n \in \mathcal{T}_\epsilon^{(n)}} p(u^n) \sum_{z^n \in \mathcal{T}_\epsilon^{(n)}(Z|U)} \exp(-2^{n(S-I(V;Z|U)-\delta(\epsilon))}\delta_1^2(\epsilon)/4) \\
&\leq 2^{n \log |\mathcal{Z}|} \exp(-2^{n(S-I(V;Z|U)-\delta(\epsilon))}\delta_1^2(\epsilon)/4),
\end{aligned}
$$

which tends to zero as $n \to \infty$ if $S > I(V; Z|U) + \delta(\epsilon)$.

We are now ready to bound $H(L|\mathcal{C}, Z^n, U^n)$. Consider

$$
\begin{aligned}
H(L, E|\mathcal{C}, U^n, Z^n) &\leq 1 + \mathrm{P}\{E = 1\}H(L|\mathcal{C}, E = 1, U^n, Z^n) + \mathrm{P}\{E = 0\}H(L|\mathcal{C}, E = 0, U^n, Z^n) \\
&\leq 1 + \mathrm{P}\{E = 1\}nS + \log((1 + \delta_1(\epsilon))2^{n(S-I(V;Z|U)+\delta(\epsilon))}) \\
&\leq n(S - I(V; Z|U) + \delta'(\epsilon)).
\end{aligned}
$$

This completes the proof of the lemma.

## APPENDIX II
### EVALUATION FOR EXAMPLE

We first give an upper bound for the extended Csiszár–Körner lower bound.

*Fact*: The extended Csiszár and Körner lower bound in (6) for the channel shown in Figure 2 is upper bounded by

$$
\begin{aligned}
R_{\mathrm{CK}} \leq \min\{&I(X_1; Y_{11}) - I(X_1; Z_1) + I(V_2; Y_{12}|Q_2) - I(V_2; Z_2|Q_2), I(X_1; Y_{21}) \\
&- I(X_1; Z_1) - I(V_2; Z_2|Q_2)\}.
\end{aligned}
$$

for some $p(x_1)p(q_2, v_2)p(x_2|v_2)$.

*Proof:* From (6), we have

$$R \leq \max_{p(q)p(v|q)p(x|v)} \min\{I(V;Y_1|Q) - I(V;Z|Q), I(V;Y_2|Q) - I(V;Z|Q)\}.$$

Consider the first bound for $R_{\text{CK}}$.

$$
\begin{aligned}
I(V;Y_1|Q) - I(V;Z|Q) &= I(V;Y_{11},Y_{12}|Q) - I(V;Z_1|Q) - I(V;Z_2|Q,Z_1) \\
&\leq I(V;Y_{11},Y_{12},Z_1|Q) - I(V;Z_1|Q) - I(V;Z_2|Q,Z_1) \\
&= I(V;Y_{11},Y_{12}|Q,Z_1) - I(V;Z_2|Q,Z_1) \\
&= I(V;Y_{11}|Q,Z_1,Y_{12}) + I(V;Y_{12}|Q,Z_1) - I(V;Z_2|Q,Z_1) \\
&= I(V;Y_{11},Z_1|Q,Y_{12}) - I(V;Z_1|Q,Y_{12}) + I(V;Y_{12}|Q,Z_1) - I(V;Z_2|Q,Z_1) \\
&\overset{(a)}{=} I(V;Y_{11}|Q,Y_{12}) - I(V;Z_1|Q,Y_{12}) + I(V;Y_{12}|Q,Z_1) - I(V;Z_2|Q,Z_1) \\
&\leq I(V';Y_{11}|Q) - I(V';Z_1|Q) + I(V;Y_{12}|Q,Z_1) - I(V;Z_2|Q,Z_1).
\end{aligned}
$$

$(a)$ follows from the structure of the channel which gives the Markov condition $(Q,Y_{12},V) - Y_{11} - Z_1$. The last step follows from defining $V' = (V,Y_{12})$ and the fact that $Z_1$ is a degraded version of $Y_{11}$.

Consider now the second bound.

$$
\begin{aligned}
R_{\text{CK}} &\leq I(V;Y_2|Q) - I(V;Z|Q) \\
&= I(V;Y_{21}|Q) - I(V;Z_1|Q) - I(V;Z_2|Q,Z_1) \\
&\leq I(V';Y_{21}|Q) - I(V';Z_1|Q) - I(V;Z_2|Q,Z_1).
\end{aligned}
$$

Combining the bounds, we have

$$R_{\text{CK}} \leq \max_{p(q,v,v',x_1,x_2)} \min \{I(V';Y_{11}|Q) - I(V';Z_1|Q) + I(V;Y_{12}|Q,Z_1) - I(V;Z_2|Q,Z_1), \tag{19}$$
$$I(V';Y_{21}|Q) - I(V';Z_1|Q) - I(V;Z_2|Q,Z_1)\}$$

Now, we note that the terms $I(V';Y_{11}|Q) - I(V';Z_1|Q)$ and $I(V';Y_{21}|Q) - I(V';Z_1|Q)$ depends only on the marginal distribution $p(q,v',x_1)p(y_{21},y_{11},z_1|x_1)$. Similarly, define $Q' = (Q,Z_1)$, the terms $I(V;Y_{12}|Q') - I(V;Z_2|Q')$ and $I(V;Z_2|Q')$ depends only on the marginal distribution $p(q',v,x_2)p(y_{12},z_2|x_2)$. Therefore, we can further upper bound $R_{CK}$ by

$$R_{\text{CK}} \leq \max \min \{I(V_1;Y_{11}|Q_1) - I(V_1;Z_1|Q_1) + I(V_2;Y_{12}|Q_2) - I(V_2;Z_2|Q_2),$$
$$I(V_1;Y_{21}|Q_1) - I(V_1;Z_1|Q_1) - I(V_2;Z_2|Q_2)\},$$

where the maximum is over $p(q_1)p(v_1|q_1)p(x_1|v_1)$ and $p(q_2)p(v_2|q_2)p(x_2|v_2)$ [1]. We now further simplify this bound as follows.

$$R_{\text{CK}} \leq \max \min \{I(V_1;Y_{11}|Q_1) - I(V_1;Z_1|Q_1) + I(V_2;Y_{12}|Q_2) - I(V_2;Z_2|Q_2),$$
$$I(V_1;Y_{21}|Q_1) - I(V_1;Z_1|Q_1) - I(V_2;Z_2|Q_2)\},$$
$$\leq \max \min \{I(X_1;Y_{11}) - I(X_1;Z_1) + I(V_2;Y_{12}|Q_2) - I(V_2;Z_2|Q_2),$$
$$I(X_1;Y_{21}) - I(X_1;Z_1) - I(V_2;Z_2|Q_2)\},$$

where the maximum is now over distributions of the form $p(x_1)$ and $p(q_2)p(v_2|q_2)p(x_2|v_2)$. The last step follows from the fact that $Z_1$ is degraded with respect to both $Y_{21}$ and $Y_{11}$. ∎

Next, we evaluate this upper bound. We will make use of the entropy relationship [17]: $H(ap, 1-p, (1-a)p) = H(p, 1-p) + pH(a, 1-a)$. First consider the terms for the first channel components,

[1]To see that this bound is larger than the previous bound in (19), set $V_1 = V'$, $Q_1 = Q$, $V_2 = (V,Q')$ and $Q_2 = Q'$ in this bound to recover the previous bound

$(I(X_1; Y_{11}) - I(X_1; Z_1))$ and $(I(X_1; Y_{21}) - I(X_1; Z_1))$. Letting $\mathrm{P}\{X_1 = 0\} = \gamma$ and evaluating the individual expressions, we obtain

$$I(X_1; Y_{21}) = H(\gamma, 1 - \gamma),$$
$$I(X_1; Y_{11}) = H\left(\frac{\gamma}{2}, \frac{1}{2}, \frac{1 - \gamma}{2}\right) - 1$$
$$= \frac{1}{2}H(\gamma, 1 - \gamma),$$
$$I(X_1; Z_1) = H\left(\frac{\gamma}{6}, \frac{5}{6}, \frac{5(1 - \gamma)}{6}\right) - H\left(\frac{1}{6}, \frac{5}{6}\right)$$
$$= \frac{1}{6}H(\gamma, 1 - \gamma).$$

This gives

$$I(X_1; Y_{21}) - I(X_1; Z_1) = \frac{5}{6}H(\gamma, 1 - \gamma),$$
$$I(X_1; Y_{11}) - I(X_1; Z_1) = \frac{1}{3}H(\gamma, 1 - \gamma).$$

Note that both expressions are maximized by setting $\gamma = 1/2$, which yields

$$R_{\mathrm{CK}} \leq \min\left\{\frac{1}{3} + I(V_2; Y_{12}|Q_2) - I(V_2; Z_2|Q_2), \frac{5}{6} - I(V_2; Z_2|Q_2)\right\}. \tag{20}$$

Next, we consider the second channel component terms. Let $\alpha_i = p(q_{2i})$, $\beta_{j,i} = p(v_{2j}|q_{2i})$, $\mathrm{P}\{X_2 = 0|V_2 = v_{2j}\} = \mu_j$, and $\mathrm{P}\{V_2 = v_{2j}\} = \nu_j$, then

$$I(V_2; Z_2|Q_2) = \sum_i \alpha_i H\left(\frac{\sum_j \beta_{j,i}\mu_j}{2}, \frac{1}{2}, \frac{\sum_j \beta_{j,i}(1 - \mu_j)}{2}\right) - \sum_j \nu_j H\left(\frac{\mu_j}{2}, \frac{1}{2}, \frac{(1 - \mu_j)}{2}\right)$$
$$= \frac{1}{2}\sum_i \alpha_i H\left(\sum_j \beta_{j,i}\mu_j, \sum_j \beta_{j,i}(1 - \mu_j)\right) - \frac{1}{2}\sum_j \nu_j H\left(\mu_j, (1 - \mu_j)\right),$$

$$I(V_2; Y_{12}|Q_2) = \sum_i \alpha_i H\left(\sum_j \beta_{j,i}\mu_j, \sum_j \beta_{j,i}(1 - \mu_j)\right) - \sum_j \nu_j H\left(\mu_j, (1 - \mu_j)\right).$$

This implies that

$$I(V_2; Y_{12}|Q_2) - I(V_2; Z_2|Q_2) = \frac{1}{2}\sum_i \alpha_i H\left(\sum_j \beta_{j,i}\mu_j, \sum_j \beta_{j,i}(1 - \mu_j)\right) - \frac{1}{2}\sum_j \nu_j H\left(\mu_j, (1 - \mu_j)\right).$$

Comparing the above expressions, we see that $I(V_2; Z_2|Q_2) = 0$ implies that $I(V_2; Y_{12}|Q_2) - I(V_2; Z_2|Q_2) = 0$. This, together with (20), implies that $R_{\mathrm{CK}}$ is *strictly* less than $5/6$.

In comparison, consider the new lower bound in Corollary 1. Setting $V = X_1$ and $X_1$ and $X_2$ independent Bernoulli $1/2$, we have

$$I(X_1, X_2; Y_{11}, Y_{12}) - I(X_1, X_2; Z_1, Z_2) = I(X_1; Y_{11}) - I(X_1; Z_1) + I(X_2; Y_{12}) - I(X_2; Z_2)$$
$$= \frac{1}{3} + \frac{1}{2} = \frac{5}{6},$$
$$I(V; Y_2) - I(V; Z) = I(X_1; Y_{21}) - I(X_1; Z_1, Z_2)$$
$$= I(X_1; Y_{21}) - I(X_1; Z_1) = \frac{5}{6}.$$

Thus, $R = 5/6$ is achievable using the new scheme, which shows that the our lower bound can be strictly larger than the extended Csiszár and Körner lower bound. In fact, $R = 5/6$ is the capacity for this example since the channel is a special case of the reversely degraded broadcast channel considered in [5] and we can use the converse result therein to show that $C_S \leq 5/6$.

<div align="center">

APPENDIX III

PROOF OF THEOREM 2
</div>

Using Fourier–Motzkin elimination on the rate constraints gives the following region.

$$R_0 < I(U; Z),$$
$$R_1 < \min\{I(V_0, V_1; Y_1|U) - I(V_1; Z|V_0), I(V_0, V_2; Y_2|U) - I(V_2; Z|V_0)\},$$
$$2R_1 < I(V_0, V_1; Y_1|U) + I(V_0, V_2; Y_2|U) - I(V_1; V_2|V_0), \tag{21}$$
$$R_0 + R_1 < \min\{I(V_0, V_1; Y_1) - I(V_1; Z|V_0), I(V_0, V_2; Y_2) - I(V_2; Z|V_0)\},$$
$$R_0 + 2R_1 < I(V_0, V_1; Y_1) + I(V_0, V_2; Y_2|U) - I(V_1; V_2|V_0), \tag{22}$$
$$R_0 + 2R_1 < I(V_0, V_2; Y_2) + I(V_0, V_1; Y_1|U) - I(V_1; V_2|V_0), \tag{23}$$
$$2R_0 + 2R_1 < I(V_0, V_1; Y_1) + I(V_0, V_2; Y_2) - I(V_1; V_2|V_0), \tag{24}$$

$$R_e \leq R_1,$$
$$R_e < \min\{I(V_0, V_1; Y_1|U) - I(V_0, V_1; Z|U), I(V_0, V_2; Y_2|U) - I(V_0, V_2; Z|U)\},$$
$$2R_e < I(V_0, V_1; Y_1|U) + I(V_0, V_2; Y_2|U) - I(V_1; V_2|V_0) - 2I(V_0; Z|U), \tag{25}$$
$$R_0 + R_e < \min\{I(V_0, V_1; Y_1) - I(V_1, V_0; Z|U), I(V_0, V_2; Y_2) - I(V_2, V_0; Z|U)\},$$
$$R_0 + 2R_e < I(V_0, V_1; Y_1) + I(V_0, V_2; Y_2|U) - I(V_1; V_2|V_0) - 2I(V_0; Z|U),$$
$$R_0 + 2R_e < I(V_0, V_2; Y_2) + I(V_0, V_1; Y_1|U) - I(V_1; V_2|V_0) - 2I(V_0; Z|U),$$
$$2R_0 + 2R_e < I(V_0, V_1; Y_1) + I(V_0, V_2; Y_2) - I(V_1; V_2|V_0) - 2I(V_0; Z|U), \tag{26}$$

with the constraint of $I(V_1, V_2; Z|V_0) \leq I(V_1; Z|V_0) + I(V_2; Z|V_0) - I(V_1; V_2|V_0)$ on the set of possible probability distributions. Due to this constraint, the numbered inequalities in the above region are redundant.

We now complete the proof by using rate splitting. This is equivalent to letting $R_1 = R_1''$, $R_0 = R_0^n + R_1'$ in the above region and letting the new rates be $R_0^n$ for the common message and $R_1^n = R_1' + R_{11}''$ for the private message. Using Fourier-Motzkin to eliminate the auxiliary rates $R_1'$ and $R_1''$ then results in the following region.

$$R_0 < I(U; Z),$$
$$R_0 + R_1 < I(U; Z) + \min\{I(V_0, V_1; Y_1|U) - I(V_1; Z|V_0), I(V_0, V_2; Y_2|U) - I(V_2; Z|V_0)\},$$
$$R_0 + R_1 < \min\{I(V_0, V_1; Y_1) - I(V_1; Z|V_0), I(V_0, V_2; Y_2) - I(V_2; Z|V_0)\},$$

$$R_e \leq R_1,$$
$$R_e < \min\{I(V_0, V_1; Y_1|U) - I(V_0, V_1; Z|U), I(V_0, V_2; Y_2|U) - I(V_0, V_2; Z|U)\},$$
$$R_0 + R_e < \min\{I(V_0, V_1; Y_1) - I(V_1, V_0; Z|U), I(V_0, V_2; Y_2) - I(V_2, V_0; Z|U)\},$$
$$R_0 + 2R_e < I(V_0, V_1; Y_1) + I(V_0, V_2; Y_2|U) - I(V_1; V_2|V_0) - 2I(V_0; Z|U),$$
$$R_0 + 2R_e < I(V_0, V_2; Y_2) + I(V_0, V_1; Y_1|U) - I(V_1; V_2|V_0) - 2I(V_0; Z|U),$$

$$R_0 + R_1 + R_e < \min\{I(V_0, V_1; Y_1|U) - I(V_1; Z|V_0), I(V_0, V_2; Y_2|U) - I(V_2; Z|V_0)\}$$
$$+ \min\{I(V_0, V_1; Y_1) - I(V_1, V_0; Z|U), I(V_0, V_2; Y_2) - I(V_2, V_0; Z|U)\},$$
$$R_0 + R_1 + 2R_e < \min\{I(V_0, V_1; Y_1|U) - I(V_1; Z|V_0), I(V_0, V_2; Y_2|U) - I(V_2; Z|V_0)\}$$
$$+ I(V_0, V_1; Y_1) + I(V_0, V_2; Y_2|U) - I(V_1; V_2|V_0) - 2I(V_0; Z|U),$$

<div align="center">

22
</div>

$$R_0 + R_1 + 2R_e < \min\{I(V_0,V_1;Y_1|U) - I(V_1;Z|V_0), I(V_0,V_2;Y_2|U) - I(V_2;Z|V_0)\}$$
$$+ I(V_0,V_2;Y_2) + I(V_0,V_1;Y_1|U) - I(V_1;V_2|V_0) - 2I(V_0;Z|U).$$

Eliminating redundant inequalities then results in

$$R_0 < I(U;Z),$$
$$R_0 + R_1 < I(U;Z) + \min\{I(V_0,V_1;Y_1|U) - I(V_1;Z|V_0), I(V_0,V_2;Y_2|U) - I(V_2;Z|V_0)\},$$
$$R_0 + R_1 < \min\{I(V_0,V_1;Y_1) - I(V_1;Z|V_0), I(V_0,V_2;Y_2) - I(V_2;Z|V_0)\},$$

$$R_e \le R_1,$$
$$R_e < \min\{I(V_0,V_1;Y_1|U) - I(V_0,V_1;Z|U), I(V_0,V_2;Y_2|U) - I(V_0,V_2;Z|U)\},$$
$$R_0 + R_e < \min\{I(V_0,V_1;Y_1) - I(V_1,V_0;Z|U), I(V_0,V_2;Y_2) - I(V_2,V_0;Z|U)\},$$
$$R_0 + 2R_e < I(V_0,V_1;Y_1) + I(V_0,V_2;Y_2|U) - I(V_1;V_2|V_0) - 2I(V_0;Z|U),$$
$$R_0 + 2R_e < I(V_0,V_2;Y_2) + I(V_0,V_1;Y_1|U) - I(V_1;V_2|V_0) - 2I(V_0;Z|U),$$

$$R_0 + R_1 + 2R_e < I(V_0,V_2;Y_2|U) - I(V_2;Z|V_0) + I(V_0,V_1;Y_1)$$
$$+ I(V_0,V_2;Y_2|U) - I(V_1;V_2|V_0) - 2I(V_0;Z|U),$$
$$R_0 + R_1 + 2R_e < I(V_0,V_1;Y_1|U) - I(V_1;Z|V_0) + I(V_0,V_2;Y_2)$$
$$+ I(V_0,V_1;Y_1|U) - I(V_1;V_2|V_0) - 2I(V_0;Z|U).$$

## APPENDIX IV
### CONVERSE FOR PROPOSITION 2

The $R_1$ inequalities follow from a technique used in [8, Proposition 11]. We provide the proof here for completeness.

$$nR_1 \le \sum_i I(M_1;Y_{1i}|M_0,Y_{1,i+1}^n) + n\epsilon_n$$
$$\le \sum_i I(M_1;Y_{1i}|M_0,Y_{1,i+1}^n,Z^{i-1}) + \sum_i I(Z^{i-1};Y_{1i}|M_0,Y_{1,i+1}^n) + n\epsilon_n$$
$$\overset{(a)}{\le} \sum_i I(M_1,Y_{1,i+1}^n;Y_{1i}|M_0,Z^{i-1}) - \sum_i I(Y_{1,i+1}^n;Y_{1i}|M_0,Z^{i-1})$$
$$+ \sum_i I(Y_{1,i+1}^n;Z_i|M_0,Z^{i-1}) + n\epsilon_n$$
$$\overset{(b)}{\le} \sum_i I(X_i;Y_{1i}|M_0,Z^{i-1}) + n\epsilon_n = \sum_i I(X_i;Y_{1i}|U_i) + n\epsilon_n,$$

where $(a)$ follows by the Csiszár sum lemma; and $(b)$ follows by the assumption that $Y_1$ is less noisy than $Z$ and the data processing inequality. The other inequality involving $Y_2$ and $Z$ can be shown in a similar fashion.

We now turn to the $R_e$ inequalities. The fact that $R_e \le R_1$ is trivial. We show the other 2 inequalities. We have

$$nR_e \le I(M_1;Y_1^n|M_0) - I(M_1;Z^n|M_0) + n\epsilon_n$$
$$= \sum_{i=1}^n \big(I(M_1;Y_{1i}|M_0,Y_{1,i+1}^n) - I(M_1;Z_i|M_0,Z^{i-1})\big) + n\epsilon_n$$
$$\overset{(a)}{=} \sum_{i=1}^n \big(I(M_1,Z^{i-1};Y_{1i}|M_0,Y_{1,i+1}^n) - I(M_1,Y_{1,i+1}^n;Z_i|M_0,Z^{i-1})\big) + n\epsilon_n$$

$$\overset{(b)}{=} \sum_{i=1}^{n} \left( I(M_1; Y_{1i}|M_0, Y_{1,i+1}^n, Z^{i-1}) - I(M_1; Z_i|M_0, Z^{i-1}, Y_{1,i+1}^n) \right) + n\epsilon_n$$

$$\overset{(c)}{\leq} \sum_{i=1}^{n} \left( I(M_1, Y_{1,i+1}^n; Y_{1i}|M_0, Z^{i-1}) - I(M_1, Y_{1,i+1}^n; Z_i|M_0, Z^{i-1}) \right) + n\epsilon_n$$

$$\overset{(d)}{\leq} \sum_{i=1}^{n} \left( I(X_i; Y_{1i}|U_i) - I(X_i; Z_i|U_i) \right) + n\epsilon_n,$$

where $(a)$ and $(b)$ follow by the Csiszár sum lemma; $(c)$ follows by the less noisy assumption; $(d)$ follows by the less noisy assumption and the fact that conditioned on $(M_0, Z^{i-1})$, $(M_1, Y_{i+1}^n) \to X_i \to (Y_{1i}, Z_i)$. The second inequality involving $I(X; Y_2|U) - I(X; Z|U)$ can be proved in a similar manner. Finally, applying the independent randomization variable $Q \sim \mathcal{U}[1:n]$, i.e. uniformly distributed over $[1:n]$, and defining $U = (U_Q, Q)$, $X = X_Q$, $Y_1 = Y_{1Q}$, $Y_2 = Y_{2Q}$ and $Z = Z_Q$ then completes the proof.

APPENDIX V
PROOF OF PROPOSITION 3

In cases two to four, the codebook generation, encoding and decoding procedures are the same as Case 1, but with different rate definitions. We therefore do not repeat these steps here.

*Case 2*: Assume that $I(U_3; Z_3) - R_0 - I(U_3; Z_2|U) \geq 0$, $I(V; Y_1|U_3) - I(V; Z_2|U_3) \leq I(V; Y_1|U_3) - I(V; Z_3|U_3)$ and $R_{e3} \leq I(V; Y_1|U_3) - I(V; Z_2|U_3)$.

In this case, using the definitions of the split message and randomization rates as in case 1, we see that we can achieve $R_{e3} = I(V; Y_1|U_3) - I(V; Z_2|U_3)$ by defining $R'_{11} = I(V; Y_1|U_3) - I(V; Z_2|U_3)$ and $R''_{11} = 0$. The *equivocation rate constraints* now are

$$R_{10}^o + R_o^r > I(U_3; Z_2|U),$$
$$R_1^r + R_{11}^o > I(V; Z_2|U_3).$$

Performing Fourier-Motzkin elimination as before then yields the rate-equivocation region given in Case 2.

*Case 3*: Assume that $I(U_3; Z_3) - R_0 - I(U_3; Z_2|U) \geq 0$, $I(V; Y_1|U_3) - I(V; Z_2|U_3) \geq I(V; Y_1|U_3) - I(V; Z_3|U_3)$ In this case, since we consider only the case of $R_1 \geq I(V; Y_1|U_3) - I(V; Z_3|U_3)$, an equivocation rate of $R_{e3} = I(V; Y_1|U_3) - I(V; Z_3|U_3)$ can be achieved by setting $R'_{11} = I(V; Y_1|U_3) - I(V; Z_3|U_3)$. The constraints for this case are as follow.

*Decoding Constraints*:

$$R_0 + R_{10}^o + R_0^r + R_{10}^s < I(U_3; Z_3),$$
$$R_{10}^s + R_{10}^o + R_0^r < I(U_3; Y_1|U),$$
$$R'_{11} + R''_{11} + R_{11}^o + R_1^r < I(V; Y_1|U_3).$$

*Equivocation rate constraints*:

$$R_{10}^o + R_o^r > I(U_3; Z_2|U),$$
$$R''_{11} + R_1^r + R_{11}^o > I(V; Z_3|U_3),$$
$$R_1^r + R_{11}^o > I(V; Z_2|U_3).$$

*Greater than or equal to zero constraints*:

$$R_{10}^o, R_0^o, R'_{11}, R''_{11}, R_1^r, R_0^r \geq 0.$$

*Equality constraints*:

$$R_1 = R_{10}^o + R_{10}^s + R'_{11} + R''_{11} + R_{11}^o,$$
$$R_{e2} = R_{10}^s + R'_{11} + R''_{11},$$
$$R_{e3} = R'_{11},$$

24

$$R'_{11} = I(V; Y_1|U_3) - I(V; Z_3|U_3).$$

Performing Fourier-Motzkin elimination then results in the rate-equivocation region for Case 3.

*Case 4*: Assume that $I(U_3; Z_3) - R_0 - I(U_3; Z_2|U) \leq 0$. In this case, note that $R_{e2} \leq \min\{R_1, I(V; Y_1|U_3) - I(V; Z_2|U_3)\}$ and can be achieved using only the $V^n$ layer of codewords. We set $R^s_{10} = 0$ in this case. If $I(V; Y_1|U_3) - I(V; Z_2|U_3) \leq I(V; Y_1|U_3) - I(V; Z_3|U_3)$, then $R_{e2} = I(V; Y_1|U_3) - I(V; Z_2|U_3)$ and $R_{e3} = \min\{R_1, I(V; Y_1|U_3) - I(V; Z_3|U_3)\}$ are achievable. If $I(V; Y_1|U_3) - I(V; Z_2|U_3) \geq I(V; Y_1|U_3) - I(V; Z_3|U_3)$, then $R_{e3} = I(V; Y_1|U_3) - I(V; Z_3|U_3)$ and $R_{e2} = \min\{R_1, I(V; Y_1|U_3) - I(V; Z_2|U_3)\}$ are achievable.

## APPENDIX VI
### PROOF OF PROPOSITION 4

As in [8], we establish bounds for the channel from $X$ to $(Y_1, Z_2)$ and for the channel from $X$ to $(Y_1, Z_3)$. *The $X$ to $(Y_1, Z_2)$ bound*: We first prove bounds on $R_0$ and $R_1$. Define the auxiliary random variables $U_i = (M_0, Y_1^{i-1})$, $U_{3i} = (M_0, Y_1^{i-1}, Z_{3,i+1}^n)$, and $V_i = (M_1, M_0, Z_{3,i+1}^n, Y_1^{i-1})$ for $i = 1, 2, \ldots, n$. Then, following the steps of the converse proof in [16], it is straightforward to show that

$$R_0 \leq \frac{1}{n} \sum_{i=1}^n I(U_i; Z_{2i}) + \epsilon_n,$$

$$R_1 \leq \frac{1}{n} \sum_{i=1}^n (I(V_i; Y_{1i}|U_i)) + \epsilon_n,$$

where $\epsilon_n \to 0$ with $n$.

To bound $R_{e2}$, first consider

$$
\begin{aligned}
H(M_1|Z_2^n) &\overset{(a)}{\leq} H(M_1|Z_2^n, M_0) + n\epsilon_n \\
&\overset{(b)}{=} H(M_1) - I(M_1; Z_2^n|M_0) + n\epsilon_n \\
&\overset{(c)}{\leq} I(M_1; Y_1^n|M_0) - I(M_1; Z_2^n|M_0) + n\epsilon_n \\
&= \sum_{i=1}^n (I(M_1; Y_{1i}|M_0, Y_1^{i-1}) - I(M_1; Z_{2i}|M_0, Z_2^{i-1})) + n\epsilon_n \\
&\overset{(d)}{\leq} \sum_{i=1}^n (I(X_i; Y_{1i}|M_0, Y_1^{i-1}) - I(X_i; Z_{2i}|M_0, Z_2^{i-1})) + n\epsilon_n \\
&= \sum_{i=1}^n (I(X_i; Y_{1i}|U_i) - H(Z_{2i}|M_0, Z_2^{i-1}) + H(; Z_{2i}|M_0, Z_2^{i-1}, X_i)) + n\epsilon_n \\
&\overset{(e)}{\leq} \sum_{i=1}^n (I(X_i; Y_{1i}|U_i) - H(Z_{2i}|M_0, Z_2^{i-1}, Y^{i-1}) + H(Z_{2i}|M_0, Z_2^{i-1}, Y^{i-1}, X_i)) + n\epsilon_n \\
&\overset{(f)}{=} \sum_{i=1}^n (I(X_i; Y_{1i}|U_i) - I(X_i; Z_{2i}|U_i)) + n\epsilon_n,
\end{aligned}
$$

where $(a)$ and $(c)$ follow by Fano's inequality, $(b)$ follows by the independence of $M_1$ and $M_0$. $(d)$, $(e)$ and $(f)$ follows by degradation of the channel from $X \to Y_1 \to Z_2$, which implies $Z_2^{i-1} \to Y_1^{i-1} \to X_i \to Y_{1i} \to Z_{2i}$ by physical degradedness. For the next inequality, we use the fact that a stochastic encoder $p(x^n|M_0, M_1)$ can be treated as a *deterministic* mapping of $(M_0, M_1)$ and an independent randomization variable $W$ onto $X^n$.

$$
\begin{aligned}
nR_0 + nR_{e2} &= H(M_0) + H(M_1|Z_2^n) \\
&\overset{(a)}{\leq} H(M_0) + H(M_1|Z_2^n, M_0) + n\epsilon_n \\
&= I(M_0; Z_3^n) + H(M_1|M_0) - H(M_1|M_0) + H(M_1|Z_2^n, M_0) + n\epsilon_n
\end{aligned}
$$

25

$$= I(M_0; Z_3^n) + I(M_1; Y_1^n | M_0) - I(M_1; Z_2^n | M_0) + n\epsilon_n$$

$$\overset{(b)}{\leq} I(M_0; Z_3^n) + I(M_1, W; Y_1^n | M_0) - I(M_1, W; Z_2^n | M_0) + n\epsilon_n$$

$$\overset{(c)}{\leq} \sum_{i=1}^{n} (I(U_{3i}; Z_{3i}) + I(X_i; Y_{1i} | U_{3i})) - I(M_1, W; Z_2^n | M_0) + n\epsilon_n$$

$$\overset{(d)}{\leq} \sum_{i=1}^{n} (I(U_{3i}; Z_{3i}) + I(X_i; Y_{1i} | U_{3i})) - \sum_{i=1}^{n} H(Z_{2i} | M_0, Y_1^{i-1})$$

$$+ \sum_{i=1}^{n} H(Z_{2i} | M_1, M_0, W, Z_2^{i-1}) + n\epsilon_n$$

$$\overset{(e)}{=} \sum_{i=1}^{n} (I(U_{3i}; Z_{3i}) + I(X_i; Y_{1i} | U_{3i})) - \sum_{i=1}^{n} H(Z_{2i} | M_0, Y_1^{i-1})$$

$$+ \sum_{i=1}^{n} H(Z_{2i} | M_1, M_0, W, Y_1^{i-1}) + n\epsilon_n$$

$$= \sum_{i=1}^{n} (I(U_{3i}; Z_{3i}) + I(X_i; Y_{1i} | U_{3i})) - \sum_{i=1}^{n} I(M_1, W, M_0; Z_{2i} | M_0, Y_1^{i-1}) + n\epsilon_n$$

$$\overset{(f)}{=} \sum_{i=1}^{n} (I(U_{3i}; Z_{3i}) + I(X_i; Y_{1i} | U_{3i})) - \sum_{i=1}^{n} I(X_i; Z_{2i} | M_0, Y_1^{i-1}) + n\epsilon_n$$

$$= \sum_{i=1}^{n} (I(U_{3i}; Z_{3i}) + I(X_i; Y_{1i} | U_{3i})) - \sum_{i=1}^{n} I(X_i; Z_{2i} | U_i) + n\epsilon_n,$$

where $(a)$ follows by Fano's inequality and $H(M_0 | Z_2^n) \leq n\epsilon_n$; $(b)$ follows by degradation of the channel from $X \to (Y_1, Z_2)$; $(c)$ by Csiszár sum applied to the first two terms (see for e.g. [2]); $(d)$ follows by the fact that conditioning reduces entropy; $(e)$ follows by the Markov relation: $Z_2^{i-1} \to Y_1^{i-1} \to (M_0, M_1, W) \to Z_{2i}$; (f) follows by the fact that $X_i$ if a function of $(M_0, M_1, W)$. This chain of inequalities implies that

$$R_{e2} \leq \frac{1}{n} \sum_{i=1}^{n} (I(U_{3i}; Z_{3i}) - I(U_{3i}; Z_{2i} | U_i)) - R_0 + \frac{1}{n} \sum_{i=1}^{n} (I(X_i; Y_{1i} | U_{3i}) - I(X_i; Z_{2i} | U_{3i})) + \epsilon_n$$

$$\leq \left[ \frac{1}{n} \sum_{i=1}^{n} (I(U_{3i}; Z_{3i}) - I(U_{3i}; Z_{2i} | U_i)) - R_0 \right]^+ + \frac{1}{n} \sum_{i=1}^{n} (I(X_i; Y_{1i} | U_{3i}) - I(X_i; Z_{2i} | U_{3i})) + \epsilon_n.$$

Finally, we arrive at single letter expressions by introducing the time-sharing random variable $Q \sim \mathcal{U}[1:n]$, i.e. uniformly distributed over $[1:n]$, independent of $(M_0, M_1, X, Y_1, Z_2, Z_3, W)$, and defining $U_Q = (M_0, Y_1^{Q-1})$, $U = (U_Q, Q)$, $V_Q = (M_1, U, Z_{2,Q+1}^n)$, $Y_1 = Y_{1Q}$ and $Z_2 = Z_{2Q}$ to obtain the following bounds

$$R_0 \leq I(U; Z_2) + \epsilon_n,$$
$$R_1 \leq I(V; Y_1 | U) + \epsilon_n,$$
$$R_{e2} \leq (I(X; Y_1 | U) - I(X; Z_2 | U)) + \epsilon_n,$$
$$R_{e2} \leq [I(U_3; Z_3) - R_0 - I(U_3; Z_2 | U)]^+ + I(X; Y_1 | U_3) - I(X; Z_2 | U_3) + \epsilon_n.$$

*The $X \to (Y_1, Z_3)$ bound*: The inequalities involving $X \to (Y_1, Z_3)$ follow standard converse techniques. First, applying the proof techniques from [15], we obtain the following bounds for the rates

$$R_0 \leq \min \left\{ \frac{1}{n} \sum_{i=1}^{n} I(U_{3i}; Z_{3i}), \frac{1}{n} \sum_{i=1}^{n} I(U_{3i}; Y_{1i}) \right\} + \epsilon_n,$$

$$R_0 + R_1 \leq \frac{1}{n} \sum_{i=1}^{n} (I(V_i; Y_{1i} | U_{3i}) + I(U_{3i}; Z_{3i})) + \epsilon_n.$$

We now turn to the second secrecy bound,

$$H(M_1|Z_3^n) \leq H(M_1, M_0|Z_3^n) = H(M_1|Z_3^n, M_0) + H(M_0|Z_3^n)$$

$$\overset{(a)}{\leq} H(M_1|Z_3^n, M_0) + n\epsilon_n$$

$$\overset{(b)}{\leq} H(M_1|Z_3^n, M_0) - H(M_1|Y_1^n, M_0) + n\epsilon_n$$

$$= I(M_1; Y_1^n|M_0) - I(M_1; Z_3^n|M_0) + n\epsilon_n,$$

where $(a)$ and $(b)$ follow by Fano's inequality. Using the Csiszár sum lemma, we can obtain the following

$$H(M_1|Z_3^n) \leq \sum_{i=1}^{n}(I(M_1; Y_{1i}|M_0, Y_1^{i-1}) - I(M_1; Z_{3i}|M_0, Z_{3,i+1}^n)) + n\epsilon_n$$

$$\overset{(a)}{=} \sum_{i=1}^{n}(I(M_1, Z_{3,i+1}^n; Y_{1i}|M_0, Y_1^{i-1}) - I(M_1, Y_1^{i-1}; Z_{3i}|M_0, Z_{3,i+1}^n)) + n\epsilon_n$$

$$\overset{(b)}{=} \sum_{i=1}^{n}(I(M_1; Y_{1i}|M_0, Y_1^{i-1}, Z_{3,i+1}^n) - I(M_1; Z_{3i}|M_0, Z_{3,i+1}^n, Y_1^{i-1})) + n\epsilon_n$$

$$= \sum_{i=1}^{n}(I(V_i; Y_{1i}|U_{3i}) - I(V_i; Z_{3i}|U_{3i})) + n\epsilon_n,$$

where both $(a)$ and $(b)$ are obtained using the the Csiszár sum lemma. Applying the independent randomization variable $Q \sim \mathcal{U}[1:n]$, i.e. uniformly distributed over $[1:n]$, we obtain

$$R_0 \leq \min\{I(U_3; Z_3), I(U_3; Y_1)\} + \epsilon_n,$$

$$R_0 + R_1 \leq I(U_3; Z_3) + I(V; Y_1|U_3) + \epsilon_n,$$

$$R_{e3} \leq I(V; Y_1|U_3) - I(V; Z_3|U_3) + \epsilon_n,$$

where $U_{3Q} = (M_0, Y_1^{Q-1}, Z_{3,Q+1}^n)$, $U_3 = (U_{3Q}, Q)$, $Y_1 = Y_{1Q}$ and $Z_3 = Z_{3Q}$. This completes the proof of the outer bound.